Physiological, Physical and Technical Demands During Sided Soccer Game Formats: a Review

Authors
Filipe Manuel Clemente[1, 2, 3], Diogo V. Martinho[4], Rui Silva[3, 5], Robert Trybulski[6, 7], Javier Sánchez-Sánchez[8], Alejandro Rodríguez-Fernández[9, 10], Marco Beato[11], José Afonso[12]

Affiliations
1 Escola Superior de Desporto e Lazer, Instituto Politécnico de Viana do Castelo, Melgaco, Portugal
2 Department of Biomechanics and Sport Engineering, Gdansk University of Physical Education and Sport, 80-336 Gdańsk, Gdansk, Poland
3 None, Sport Physical Activity and Health Research & Innovation Center, 4900-347 Viana do Castelo, Viana do Castelo, Portugal
4 Faculty of Sport Science and Physical Education, University of Coimbra, Coimbra, Portugal
5 Escola Superior de Desporto e Lazer, Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal
6 Medical Department, Wojciech Korfanty Upper Silesian Academy, Katowice, Poland
7 None, Provita Żory Medical Center, Żory, Poland
8 Faculty of Education, Universidad Pontificia de Salamanca, Salamanca, Spain
9 Faculty of Physical Activity and Sports Science, University of León, Leon, Spain
10 VALFIS Research Group, Institute of Biomedicine (IBIOMED), Leon, Spain
11 School of Health and Sports Science, University of Suffolk, Ipswich, United Kingdom of Great Britain and Northern Ireland
12 Centre of Research, Education, Innovation, and Intervention in Sport (CIFI2D), Faculty of Sport, University of Porto, FADEUP, Porto, Portugal

Keywords
football, small-sided games, conditioned games, technical

received 12.03.2025
accepted after revision 18.04.2025
accepted manuscript online 18.04.2025
published online 26.05.2025

Bibliography
Int J Sports Med 2025; 46: 710–741
DOI 10.1055/a-2591-6995
ISSN 0172-4622

Georg Thieme Verlag KG, Oswald-Hesse-Straße 50, 70469 Stuttgart, Germany

Correspondence
Dr. Filipe Manuel Clemente
Escola Superior de Desporto e Lazer
Instituto Politécnico de Viana do Castelo
Melgaco
Portugal
filipe.clemente5@gmail.com

Supplementary Material is available at https://doi.org/10.1055/a-2591-6995

ABSTRACT

This meta-analysis aimed to compare the acute physiological, physical, and technical demands in soccer players during different sided game formats (1v1 to 10v10). This review included studies on soccer players with a competitive or developmental level, focusing on games with at least one comparison of sided formats. Outcomes assessed included physiological parameters (e.g., heart rate and blood lactate levels), physical demands (e.g., distance and accelerations), and technical actions (e.g., passes). The risk of bias assessment tool for nonrandomized studies of interventions (RoBANS 2) and (Grading of Recommendations, Assessment, Development and Evaluation) were used to evaluate the risk of bias and the certainty of evidence. The search across PubMed, Scopus, and Web of Science identified 2,545 records, of which 72 studies were included. Meta-analyses found that 2v2 and 3v3 formats were more physiologically intense, showing higher perceived exertion and blood lactate levels, with 3v3 also having higher heart rates. The 4v4 and 3v3 formats resulted in greater physical demands compared to the 2v2 format, with increased distances covered at various speeds, although differences were minimal beyond the 4v4 format. Smaller formats promoted ball possession, while the 3v3 format resulted in more successful shots, dribbles, and passes. In conclusion, small-sided games (< 3v3) were more physiologically demanding, mid-sized formats (> 4v4) increased locomotor demands, and smaller formats improved technical skills, although the findings should be interpreted cautiously due to study limitations.

Introduction

Sided games are popular soccer-based training drills that adjust the complexity of the game by reducing the number of players and/or the size of the playing field [1]. To achieve the desired acute responses, coaches manipulate the task constraining variables of these drills while preserving some of the core dynamics of the original game [2]. One of the main reasons that sided games are so popular is their ability to maintain key game dynamics, even when some components are altered [3]. This allows players to remain highly engaged in achieving specific tactical and technical goals [4]. Furthermore, the modifications to game formats make these drills both challenging and physiologically intense, enabling them to serve technical and tactical developments and physical conditioning (particularly aerobic endurance as well as neuromuscular components due to accelerations and decelerations [5]).

Among the various task constraints that coaches can manipulate, the format of play (i.e., numerical relationships between teams) has been one of the most studied constraints [6]. Changing the format of play strongly influences the individual player's participation, which in turn impacts the overall demands of the game and, consequently, the physical and physiological efforts required [7]. Additionally, while tactical behavior is expected to change across different formats of play, one of the key outcomes is the effect on the individual's participation and the number of actions a player can engage in [8, 9]. These aspects – whether physiological, physical, or technical – can be directly targeted by simply adjusting the sided game format.

Building on the growing evidence surrounding sided games, various systematic reviews [1, 6, 10] on the topic have suggested that smaller formats (e.g., 1v1 to 4v4) tend to elicit greater physiological responses, such as heart rates (HRs), rating of perceived exertion (RPE), and blood lactate (BLa) concentrations. However, evidence regarding physical demands (e.g., locomotor demands as the total distance covered, distances at different speeds, accelerations, and decelerations) is less consistent [7]. This inconsistency may be influenced by the interaction between a format and a pitch size. Even when maintaining a consistent player-to-area ratio (the number of players involved in the game divided by the area of the field, excluding goalkeepers), increasing the format size naturally results in a larger playing field, allowing players to cover greater distances at higher speeds [11]. Between- and within-player variability can undermine these trends, leading to conflicting findings [12]. Regarding technical actions, studies suggest that smaller formats generally lead to a higher frequency of individuals' actions (e.g., passes, dribbles, and shots), although evidence regarding the accuracy of these actions is inconclusive [9].

While these trends are highlighted in published reviews [1, 6, 7, 9, 10], two key gaps remain to be addressed: (i) no meta-analytical comparisons have been conducted to statistically assess the magnitude of differences in acute demands elicited by various sided game formats, which would enable the clear identification of distinct categories of sided games based on the similarity of their stimuli and (ii) the variability in comparisons among individual studies makes it difficult to provide a clear understanding of the specific parameters of sided games that should be implemented based on targeted outcomes. Although a previous classification proposed by Owen et al. [13] has categorized sided games as small-sided games (SSGs: 4v4), medium-sided games (5v5 to 7v7), and large-sided games (LSGs: 8v8 to 10v10), this classification may have been somewhat arbitrary, lacking sufficient evidence to identify clear thresholds that define clusters of similar games based on comparable measures.

Based on this, our aim was to systematically review and meta-analyse the acute physiological, physical, and technical demands elicited by different sided game formats in male and female (youth and adult) soccer players.

Materials and Methods

We conducted our systematic review in alignment with the guidelines specified in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) Statement [14]. The protocol for this systematic review was made publicly available on the Open Science Framework (project: osf.io/qza3h; registration: [10.17605/OSF.IO/5XZUQ]) on December 09, 2024.

Eligibility criteria

The eligibility criteria were based on the PI/ECOS framework, which considers participants, intervention/exposure, comparator, outcomes, and study design, as outlined in ▶ **Table 1**. Eligible studies included original research articles from peer-reviewed journals, with no restrictions on the publication year [15] or article language.

Information sources

We conducted a thorough search for relevant studies across multiple databases, including PubMed, Scopus, and Web of Science (Core Collection), on December 10, 2024, after registering the review protocol. To expand our search, we also manually examined the reference lists of the included studies and utilized snowball citation tracking through the Web of Science database. In order to strengthen the review's validity, we sought input from two leading experts in soccer, as recognized by Expertscape (https://expertscape.com/ex/soccer). Furthermore, all studies incorporated into the review were carefully scrutinized for potential errors or retractions.

Information sources

The search was conducted using Boolean operators "AND" and "OR," with the intentional decision to avoid applying filters or restrictions based on the date, language, or study design. This approach was chosen to maximize the chances of identifying relevant studies. The search strategy used is outlined below:

[Title/Abstract] soccer OR football *

AND

[Title/Abstract] "small-sided games" OR "medium-sided games" OR "large-sided games" OR "sided games" OR SSG OR "conditioned games" OR "constrained games"

The complete search strategy for each database is provided in Supplementary Material S1 (available in the online version only).

Selection process

In the first phase of the search process, two authors (FMC and DM) independently screened the titles and abstracts of the retrieved studies. These abstracts were assessed against the relevant eligibility

▶ **Table 1** Eligibility criteria in this systematic review and meta-analysis.

	Inclusion criteria	Exclusion criteria
Population	Boys/men or girls/women players enrolled in soccer, with the condition that they were not reported as injured or unwell. Players were required to be actively participating in soccer training and competing at or higher than tier 2 in the Participant Classification Framework[a] [16], regardless of their age.	Players competing in parasports or those who were injured or unwell were excluded. Additionally, players participating in other football codes (e.g., futsal, beach football, indoor football, Australian football, Gaelic football, rugby football, or international rules football) or other sports were excluded.
Intervention or/and exposure	Balanced sided game formats (i.e., numerically balanced without the use of floaters or numerical disadvantages) were eligible if compared with a different format while maintaining consistent rule conditions. Alterations in field dimensions were permitted. Additionally, changes in the regimen (e.g., duration per set and/or number of repetitions) were allowed, provided that they reflected regular adjustments made to accommodate physiological demands. To be included, each format had to be played at least twice by the same players, either within the same session or across different sessions.	Exclusion criteria included formats performed only once (which could increase variability) and studies that modified task rules or objectives compared to the control group (e.g., altering the scoring method). Numerically unbalanced formats (e.g., 3v2 or 3v3 + 2) were also excluded from the analysis.
Comparator	Sided games other than the one tested in the intervention were eligible if they were numerically balanced. The comparator had to follow the same task rules (e.g., scoring methods and the number of ball touches allowed) as the tested sided game. Each format had to be repeated at least twice, either within or across sessions, under conditions comparable to the other sided game.	Comparators were excluded if they did not maintain consistent experimental conditions (e.g., task rules and objectives) or if the sided game was performed only once. Comparisons involving numerically unbalanced formats were also excluded.
Outcomes	At least one measure from the following dimensions: (i) physiological responses (e.g., HR, BLa, or RPE), (ii) physical demands (e.g., total distance, distances covered at different speed thresholds, and acceleration/deceleration), and (iii) technical actions (e.g., passes, receptions, and shots).	Tactical measures, collective behavior measures, and psychological and sociological measures were excluded. Additionally, indicators of well-being or readiness, as well as studies examining the effects of muscle fatigue, were not considered.
Study design	Counterbalanced designs were eligible.	Studies with non-counterbalanced designs were excluded.

Abbreviations: BLa: blood lactate; HR: heart rate; RPE: rating of perceived exertion.

[a]The competitive level was classified based on the Participant Classification Framework [16] where Tier 2 refers to the trained/developmental level, Tier 3 refers to the highly trained/national level, Tier 4 refers to the elite/international level, and Tier 5 refers to world class.

criteria. During the second phase, the full-text versions of the studies retained from the first phase were separately reviewed by the same two authors. In cases of disagreement at either phase, the authors engaged in further discussion, and if they could not reach a consensus, a third author (RMS) was consulted to make a final decision. To manage records efficiently and remove duplicates, a combination of manual and automated methods was used, supported using EndNot software (version 21.5, Clarivate Analytics, Philadelphia, PA).

Data collection process

The lead author (FMC) initially conducted the data extraction process, which was then carefully reviewed for accuracy and completeness by two additional authors (DM and RMS). To support this process, a dedicated Microsoft Excel spreadsheet (Microsoft, USA) was developed to capture all relevant data. A sample excerpt from this datasheet is provided in Supplementary Material S2 (available in the online version only). To address issues of missing data from the full-text versions of studies, the lead author (FMC) reached out to the corresponding author via email and ResearchGate to request the missing information. If no response was received within 2 weeks, the data from those studies were excluded from both the review and the meta-analysis.

Data items

The following general details were extracted from each included study: (i) sample size; (ii) age, sex, and competitive level of the players, as classified according to the Participant Classification Framework [16]; (iii) the sided game format; (iv) the field dimensions and player-to-area ratio; (v) specific task rules and task objectives; and (iv) the number of sessions and repetitions performed per sided game and the total duration per each format. We also collected information on contextual factors, including the phase of the season, the time of day each testing session occurred, and the randomization process applied to the sequence of sided games.

For physiological measures, the following variables were extracted, including but not limited to: (i) cardiovascular responses (e.g., mean and peak HRs), (ii) blood lactate levels, and (iii) RPE. For physical demands, the following variables were extracted, including but not limited to: (i) total distance covered, (ii) distance covered across different speed thresholds, (iii) the number of accelerations and decelerations at varying intensity levels, and (iv) mechanical workload metrics (physical effort and energy expended by the body in terms of movement and force production) derived from inertial measurement units. In terms of technical performance, the following variables were extracted, including but not limited to: (i) total and accuracy-adjusted number of passes, (ii) total and accuracy-adjusted number of receptions, (iii) total and accuracy-adjusted

number of shots, (iv) total and accuracy-adjusted number of dribbles, and (v) ball possessions and lost balls.

When data were presented graphically, a specialized software tool was utilized to extract the mean and standard deviation (SD; http://www.getdata-graph-digitizer.com). This software is known for its precision and reliability in extracting these values [17].

Risk of bias assessment

The revised risk of bias assessment tool for nonrandomized studies of interventions (RoBANS 2) was used to assess the risk of bias (RoB) in the included studies. The revised version demonstrated acceptable feasibility, fair-to-moderate reliability, and construct validity [18]. The tool evaluates the RoB across the following domains: (i) comparability of the target group, (ii) target group selection, (iii) confounders, (iv) measurement of intervention/exposure, (v) blinding of assessors, (vi) outcome assessment, (vii) incomplete outcome data, and (viii) selective outcome reporting. The global RoB was classified as unclear, high, or low, and this classification aimed to support the interpretation of the results and help determine the overall strength of evidence. Two authors (JA and RMS) independently applied this scale to assess the study quality. Any discrepancies in ratings were resolved through discussion, with a final decision made by a third author (FMC).

Summary measures, synthesis of results, and publication bias

Meta-analysis typically requires data from at least two studies to produce valid results [19]. However, due to the frequent presence of small sample sizes in sports science research [20], especially in studies on sided games [21], this meta-analysis only included data from studies that reported both means and SDs for comparisons involving at least three study groups across different sided game formats of a specific variable. To ensure consistency in data analysis and reporting, only outcomes that reported means and SDs in at least three studies were included.

To calculate effect sizes (ESs; Hedge's *g*) for each outcome, the relevant means and SDs were used. If the means and SDs were not directly available, they were estimated from 95 % confidence intervals (CIs) or standard errors of the mean using Cochrane's RevMan calculator. An integrative method was adopted to combine multiple ESs from the same study [22].

A random-effects model was applied to account for the variability among studies, which could impact the SSG outcomes [23, 24]. ESs were presented with 95 % CIs. The ES values were classified according to the following scale: < 0.2 (trivial), 0.2–0.6 (small), 0.6–1.2 (moderate), 1.2–2.0 (large), 2.0–4.0 (very large), and > 4.0 (extremely large) [25]. The impact of heterogeneity was assessed using the I^2 statistic, with values of < 25, 25–75, and > 75 % indicating low, moderate, and high heterogeneity, respectively [26]. The 95 % CIs were assessed to describe the precision of the combined effect. The risk of publication bias was evaluated with the extended Egger's test only when at least 10 studies were available per comparison [26]. Additionally, potential contributors to heterogeneity were investigated including the competitive level, training regimen and sided game field dimensions by segregating the meta-analyses based on each of these variables [27]. To address potential publication bias, a sensitivity analysis was performed using the trim and fill method [28]. All statistical analyses were conducted using comprehensive meta-analysis software (version 2; Biostat, Englewood, NJ, USA), with statistical significance set at $p < 0.05$.

To examine the effect of pitch size, we included studies that maintained a similar player-to-area ratio. Based on the median value of all studies (i.e., 100 m^2), two groups were created: ≤ 100 and > 100 m^2. Comparisons between these two groups regarding different game formats were conducted only for variables that had at least three studies in each group.

Certainty assessment

The assessment focused on the five criteria outlined in the GRADE framework [29, 30] which include the RoB, indirectness, inconsistency, the risk of publication bias [27], and imprecision (all factors that can downgrade the certainty of evidence). These factors were used to determine the quality of evidence related to load and performance outcomes, categorizing it as high, moderate, low, or very low. Since all the studies were non-randomized, they initially received a low-quality rating. However, the quality could be increased if there were substantial ESs, effective control over potential confounders, and evidence of a dose–response relationship. Upgrades were only made if there were no grounds for downgrading, in accordance with the GRADE guidelines [31–34].

To evaluate the confidence in evidence, we established a comprehensive set of criteria. Initially, we assessed the RoB in the included studies. If a moderate RoB was identified, the evidence was downgraded by one level. In cases where a high RoB was observed, a two-level downgrade was applied. We considered the evidence to have low indirectness by default, as the study populations, exposures, and outcomes met the eligibility criteria for direct relevance. Regarding the publication bias, we chose not to evaluate this risk, as each analysis included fewer than 10 studies [27]. For inconsistency, heterogeneity was evaluated based on the 95 % CI [27]. Finally, imprecision was assessed by considering the sample size and the clarity of the effect. Evidence was downgraded by one level if the sample size was smaller than 800 participants (fewer than 400 per group [35] or if the effect direction was unclear (e.g., when the 95 % CI crossed zero [35]. Evidence was downgraded by two levels for imprecision in the case where both factors were present.

Results

Study identification and selection

Database searches identified a total of 2,545 records. After removing duplicates ($n = 1,230$), 1,315 studies remained for screening based on their titles and abstracts. Of these, 1,197 studies were excluded, resulting in 118 studies selected for the full-text review. During the full-text assessment, 47 studies were deemed ineligible based on the predefined criteria (see Supplementary Material S2, available in the online version only). Consequently, 71 studies were included in the review. Additionally, one more study was identified through alternative methods, such as expert recommendations. This brought the total number of studies included in this review to 72.

RoB assessment

The RoBAN 2 tool was used to assess the RoB, revealing that while the RoB is low for comparability of the target group (98.6 % of studies show low risk), target group selection (100 % of studies show low risk), measurement of intervention/exposure (98.6 % of studies show low risk), outcome assessment (100 % of studies show low risk), and blinding of assessors (79.2 % of studies), there is an unclear RoB in incomplete outcome data (59.7 % of studies) and selective outcome reporting (90.3 % of studies). Additionally, there is a tendency toward a high RoB in confounding (56.9 % of studies show high risk). ▶ **Table 2** presents the specific RoB for each study.

Study characteristics

▶ **Table 3** exhibits the characteristics of the included studies. A total of 72 studies included 1,767 participants, with a median of 18 participants per study. The number of participants ranged from a minimum of 8 [36] to a maximum of 197 [54]. In terms of competitive levels, Tier 2 (trained/developmental) was the most frequent, with 43 studies, followed by Tier 3 (highly trained, national) with 20 studies, Tier 4 (elite/international) with 8 studies, and Tier 5 (world class) with 1 study [98]. Regarding sex, men were the most commonly analyzed ($n = 44$), with three studies [88, 94, 101] exclusively focused on women and two studies [50, 51] including both men and women. Twenty-three studies did not report the sex of the participants. In terms of age, 29 studies had an average participants' age of greater than 20, while 42 studies had an average age of 18 years or younger.

Regarding the dimensions of analysis, 48 studies included psychophysiological measures, followed by 42 studies that analyzed physical demands. Technical elements were assessed in 27 studies. Nine studies integrated both physiological and physical measures, 16 studies integrated both physiological and technical measures, and 7 studies integrated both physical and technical measures.

▶ **Table 4** presents the characteristics of the sided games. The 1v1 format was analyzed in 7 studies, 2v2 in 28 studies, 3v3 in 45 studies, 4v4 in 45 studies, 5v5 in 25 studies, 6v6 in 19 studies, 7v7 in 14 studies, 8v8 in 10 studies, 9v9 in 5 studies, and 10v10 in 5 studies. Each format may have included or excluded a goalkeeper. For example, 33 studies involved goalkeepers, while 18 studies focused on sided games with ball possession as the main task objective, 20 studies analyzed sided games with small goals, and 3 studies employed the stop-ball rule as the primary task objective.

Results of individual studies and meta-analysis

Physiological outcomes

▶ **Table 5** summarizes the meta-analysis parameters of different formats of sided games in relation to physiological outcomes, including RPE, maximal HRs, average HRs, and BLa levels. The results indicate that the RPE was higher in smaller game formats. All the comparisons performed were statistically significant, with the overall Hedge's *g*-values ranging from − 1.917 (large ES) to − 0.524 (small ES). Heterogeneity levels were categorized as follows: low in three comparisons (2v2 vs. 5v5, 4v4 vs. 5v5, and 5v5 vs. 6v6), moderate in four comparisons (2v2 vs. 3v3, 3v3 vs. 4v4, 3v3 vs. 6v6, and 4v4 vs. 6v6), and high in two comparisons (2v2 vs. 4v4 and 3v3 vs. 5v5). Additionally, the risk of publication bias was identified in three comparisons: 2v2 vs. 3v3 (Egger's intercept = − 6.251; $p < 0.05$), 2v2 vs. 4v4 (Egger's intercept = − 9.254; $p < 0.01$), and 3v3 vs. 4v4 (Egger's intercept = − 7.769; $p < 0.01$).

The maximal HR was significantly higher in the 3v3 format compared to the 4v4 format (Hedge's $g = -0.467$), and also in the 4v4 format when contrasted with the 5v5 format (Hedge's $g = -0.365$) and 6v6 format (Hedge's $g = -0.856$). Furthermore, the maximal HR was significantly higher in the 5v5 format compared to the 6v6 game format (Hedge's $g = -0.355$), with low heterogeneity ($I^2 = 20.8\,\%$).

The average HR was largely and moderately greater in the comparisons of 2v2 vs. 4v4 and 3v3 vs. 4v4, respectively, with high heterogeneity observed (2v2 vs. 4v4: $I^2 = 92.6\,\%$; 3v3 vs. 4v4: $I^2 = 79.1\,\%$). In the first comparison, no trimmed studies were identified.

BLa levels were, on average, higher in smaller game formats. The magnitude of differences was moderate when comparing 2v2 to 3v3, very large when contrasting 2v2 with 4v4, and large when comparing 3v3 to 4v4. Heterogeneity values ranged from 34 % to 88 %, with the publication bias identified in the comparisons of 2v2 vs. 3v3 (Egger's intercept = − 9.99; $p < 0.01$), 2v2 vs. 4v4 (Egger's intercept = − 12.729; $p < 0.01$) and 3v3 vs. 4v4 (Egger's intercept = − 14.856; $p < 0.01$). In these comparisons, trimmed studies were identified but the adjusted values did not change considerably.

By accessing the public repository of the review at https://osf.io/qza3h/, readers can obtain the raw data included in the meta-analysis for all assessed measures, as well as the forest plots for physiological, physical, and technical outcomes.

Physical outcomes

▶ **Table 6** highlights the statistical analysis of sided game physical demands across various game formats. The total distance covered was significantly higher in larger game formats. The differences were very large when comparing 2v2 with 3v3 (Hedge's $g = 2.635$) and 4v4 (Hedge's $g = 4.095$). In contrast between 3v3 and 4v4, the differences were large (Hedge's $g = 1.299$). Heterogeneity in these comparisons was classified as high, and the publication bias was observed across all three meta-analyses. No trimmed studies were identified. Additionally, the total distance covered was significantly higher in the 5v5 format compared to the 10v10 format (Hedge's $g = 1.207$) and also higher in the 6v6 format compared to the 9v9 format (Hedge's $g = 1.483$) and 10v10 format (Hedge's $g = 0.743$).

The distance covered in the range of 0–6.9 km h^{-1} was higher in 4v4 compared to both 2v2 (Hedge's $g = 2.130$) and 3v3 (Hedge's $g = 1.289$). Heterogeneity values in these comparisons were high. Additionally, the distance covered in the range of 0–6.9 km h^{-1} was significantly greater in 5v5 than in 3v3 (Hedge's $g = 0.546$), with moderate heterogeneity across studies.

For distances covered in the ranges of 7–13 km h^{-1} and 13–17.8 km h^{-1}, the results indicated that 4v4 showed significantly greater distances compared to 2v2 and 3v3. The magnitude of differences varied from small (Hedge's $g = 0.574$) to very large (Hedge's $g = 2.525$). No risk of publication bias or trimmed studies was identified for the 7–13 km/h range (2v2 vs. 4v4 comparison).

The distance covered at a speed greater than 18 km h^{-1} favored the larger game formats. The differences between 2v2 and 3v3 (Hedge's $g = 1.031$) and between 2v2 and 4v4 (Hedge's $g = 1.314$)

▶ **Table 2** Risk of bias assessment using the revised version of the risk of bias assessment tool for nonrandomized studies (RoBANS 2).

Study	Comparability of the target group	Target group selection	Confounders	Measurement of intervention/exposure	Blinding of assessors	Outcome assessment	Incomplete outcome data	Selective outcome reporting
Aasgaard et al. [36]	Low	Low	Low	Low	Low	Low	High	Unclear
Abrantes et al. [37]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Ade et al. [38]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Aguiar et al. [39]	Low	Low	High	Low	Low	Low	High	Unclear
Alashti et al. [40]	Low	Low	Low	Low	Low	Low	Low	Unclear
Alcântara et al. [41]	Low	Low	Low	Low	Low	Low	High	Unclear
Almeida et al. [42]	High	Low	Low	Low	High	Low	Unclear	Unclear
Arslan et al. [43]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Aşçı et al. [44]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Asian-Clemente et al. [45]	Low	Low	Low	Low	Low	Low	Low	Unclear
Asín-Izquierdo et al. [46]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Beato et al. [47]	Low	Low	Low	Low	Low	Low	High	Unclear
Beato et al. [48]	Low	Low	High	Low	Low	Low	Low	Unclear
Beenham et al. [49]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Bergmann et al. [50]	Low	Low	High	Low	High	Low	Unclear	High
Bergmann et al. [51]	Low	Low	High	Low	Unclear	Low	Unclear	High
Brandes et al. [52]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Bravo-Sánchez et al. [53]	Low	Low	Low	Low	Low	Low	Low	Unclear
Brito et al. [54]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Casamichana et al. [55]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Castellano et al. [56]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Castellano et al. [57]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Clemente [58]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Clemente et al. [59]	Low	Low	High	Low	Low	Low	Low	Unclear
Clemente et al. [60]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Clemente et al. [61]	Low	Low	High	Low	Low	Low	Low	Unclear
Clemente et al. [62]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Clemente et al. [63]	Low	Low	High	Low	Low	Low	Low	Unclear
Costa et al. [64]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Da Silva et al. [65]	Low	Low	High	Low	Low	Low	Low	Unclear
Dalen et al. [66]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Dellal et al. [67]	Low	Low	High	Low	Low	Low	High	Unclear
Dellal et al. [68]	Low	Low	Low	Low	High	Low	Unclear	High
Dellal et al. [69]	Low	Low	Low	Low	Unclear	Low	Unclear	High
Dellal et al. [70]	Low	Low	High	Low	Low	Low	Low	Unclear
Dellal et al. [71]	Low	Low	High	Low	Low	Low	Unclear	High
Dimitriadis et al. [72]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Edis et al. [73]	Low	Low	High	Low	High	Low	Low	Unclear
Evangelos et al. [74]	Low	Low	High	Low	High	Low	High	Unclear
Farhani et al. [75]	Low	Low	Low	Low	High	Low	Unclear	Unclear
Febré et al. [76]	Low	Low	Low	Low	High	Low	Low	Unclear
García-Angulo et al. [77]	Low	Low	High	Low	High	Low	Unclear	Unclear
Giménez et al. [78]	Low	Low	High	Low	Low	Low	High	Unclear
Halouani et al. [79]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Harrison et al. [80]	Low	Low	Low	Low	High	Low	Unclear	Unclear
Hill-Haas et al. [81]	Low	Low	High	Low	Low	Low	High	Unclear
Jastrzębski et al. [82]	Low	Low	High	Low	Low	Low	High	Unclear
Köklü et al. [83]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Köklü [84]	Low	Low	Low	Low	Low	Low	Unclear	Unclear

▶ **Table 2** Continued. nt tool for nonrandomized studies (RoBANS 2).

Köklü et al. [85]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Köklü et al. [86]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Köklü et al. [87]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Köklü et al. [88]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Köklü et al. [89]	Low	Low	Low	Low	Low	Low	High	Unclear
Little et al. [90]	Low	Low	High	High	High	Low	High	Unclear
Lex et al. [91]	Low	Low	High	Low	Low	Low	Low	Unclear
López-Fernández et al. [92]	Low	Low	High	Low	Low	Low	Low	Unclear
Makar et al. [93]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Mara et al. [94]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Martin-Garcia et al. [95]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Owen et al. [96]	Low	Low	High	Low	High	Low	High	Unclear
Owen et al. [13]	Low	Low	Low	Low	Low	Low	High	Unclear
Papadopoulos et al. [97]	Low	Low	Low	Low	Low	Low	High	Unclear
Pinheiro et al. [98]	Low	Low	High	Low	Low	Low	High	Unclear
Rampinini et al. [99]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Rebelo et al. [100]	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Savolainen et al. [101]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Silva et al. [102]	Low	Low	High	Low	Low	Low	Unclear	Unclear
Wang et al. [103]	Low	Low	High	Low	High	Low	Low	Unclear
Yilmaz et al. [104]	Low	Low	Low	Low	High	Low	Unclear	Unclear
Younesi et al. [105]	Low	Low	High	Low	Low	Low	Unclear	High
Younesi et al. [106]	Low	Low	High	Low	Low	Low	Unclear	High

were classified as moderate and large, respectively. Both comparisons exhibited high heterogeneity. In the comparison of 3v3 vs. 4v4 for the distance covered > 18 km h^{-1}, the results favored the 4v4 format, with a small magnitude of difference between the game formats (Hedge's $g = 0.454$).

No differences between formats were noted for the following variables: the distance covered greater than 19 km h^{-1}, the number of accelerations, and the number of decelerations. However, the number of sprints was higher in the 4v4 format compared to the 2v2 format, favoring the larger soccer format (Hedge's $g = 0.418$), with low heterogeneity observed. It is important to note that the comparisons of the number of accelerations, decelerations, and sprints included different thresholds used across the studies.

Technical outcomes

▶ **Table 7** describes the meta-analysis data comparing technical variables across different sided games formats. The number of passes completed was significantly higher in 3v3 compared to 2v2 (Hedge's $g = 1.175$), with high heterogeneity observed ($I^2 = 81.5\%$). In contrast, the number of dribbles performed was very significantly higher in 3v3 compared to 4v4 (Hedge's $g = -3.939$). Heterogeneity in this comparison was also high ($I^2 = 96.3\%$).

The number of successful shots (in case of games with small goals or regular goals) was higher in 3v3 compared to both 5v5 (Hedge's $g = -0.763$) and 6v6 (Hedge's $g = -1.060$). The heterogeneity was moderate in both comparisons (3v3 vs. 5v5: $I^2 = 39.6\%$; 3v3 vs. 6v6: $I^2 = 70.1\%$). Ball possession was significantly higher in 4v4 compared to 2v2 (Hedge's $g = -1.948$) and 3v3 (Hedge's $g = 1.593$; ▶ **Fig. 1**).

Comparing play formats across different player-to-area ratios

Sub-group analyses were conducted for three variables: the rate of perceived exertion, maximal HR, and total distance covered, across three game formats (2v2 vs. 3v3, 2v2 vs. 4v4, and 3v3 vs. 4v4). ▶ **Fig. 2** presents the sub-group analyses for the rate of perceived exertion (panel A), maximal HR (panel B), and total distance covered (panel C). The overall effect of the rate of perceived exertion was significantly higher in the 2v2 format (Hedge's $g = -0.985$, $p < 0.05$). For studies with a player-to-area ratio below or equal to the median of all studies, the effect was non-significant (Hedge's $g = -0.624$, $p = 0.206$). In contrast, studies with a player-to-area ratio of greater than 100 m^2 yielded a significant effect (Hedge's $g = -1.475$, $p < 0.05$). For the maximal HR, the comparisons across both groups showed non-significant differences. In terms of the total distance covered, the 3v3 format was favored over the 2v2 format, demonstrating very large effects for a player-to-area ratio of ≤ 100 m^2 (Hedge's $g = 3.963$, $p < 0.05$). However, for a player-to-area ratio of > 100 m^2, no significant differences were found (Hedge's $g = -0.210$, $p = 0.88$).

When comparing the 2v2 and 4v4 formats, the rate of perceived exertion was higher in the 2v2 format (Hedge's $g = 1.958$, $p < 0.05$). Sub-group analyses indicated moderate effects for player-to-area ratios of ≤ 100 m^2 (Hedge's $g = -1.107$, $p < 0.05$) and large effects for those > 100 m^2 (Hedge's $g = -1.958$, $p < 0.05$). The differences in the maximal HR between game formats were non-significant. In terms of the total distance covered, the 4v4 format produced significantly higher results for player-to-area ratios of ≤ 100 m^2 (Hedge's $g = 4.512$, $p < 0.05$), while negligible differences were

▶ **Table 3** Characteristics of the included studies.

Study	Competitive level	Sex	*N*	Age (y)	Physiological measures extracted	Physical demands extracted	Technical actions extracted
Aasgaard et al. [36]	Tier 2	n.d.	8	23.6 ± 3.3	HRpeak (%), HRave (%), RPE (A.U.)	TD (m), workload (m/min), speed max (km/h), speed average (km/h), and TD at different speed thresholds	n.a.
Abrantes et al. [37]	Tier 2	Men	16	15.8 ± 0.5	HR zones 1–5 (min) and RPE (A.U.)	n.a.	Pass, receive, dribble, shot, tackle, and interception
Ade et al. [38]	Tier 3	Men	16	17.0 ± 1.0	HRave (%), HRpeak (%), BLa (mmol/L), and RPE (A.U.)	TD (m), TD at different speed thresholds, acceleration and deceleration distances at 2–3 and > 3 m/s^2	n.a.
Aguiar et al. [39]	Tier 3	Men	10	18.0 ± 0.7	HRmax (%) and HR zones 1–5 (min)	TD (m), TD at different speed thresholds, and accumulated load (A.U.)	n.a.
Alashti et al. [40]	Tier 2	Men	24	17.1 ± 1.1	HR (bpm), HRmax (%) and BLa (mmol/L)	n.a.	n.a.
Alcântara et al. [41]	Tier 2	Men	11	13.0 ± 0.8	RPE (A.U.)	TD (m) and TD at different speed thresholds	Successful passes, unsuccessful passes, contacts with the ball, ball involvement, goals scored, shots on the target, and unsuccessful shots
Almeida et al. [42]	Tier 2	Men	28	12.9 ± 0.6	n.a.	n.a.	Ball possession, ball touches, passes, shots and goals
Arslan et al. [43]	Tier 2	n.d.	16	16.9 ± 0.3	HR (bpm), HRmax (%), BLa (mmol/L), and RPE (A.U.)	TD (m) and TD at different speed thresholds	n.a.
Aşçı et al. [44]	Tier 3	Men	22	17.4 ± 0.9	HR (bpm), HRmax (%), and HR zones	n.a.	n.a.
Asian-Clemente et al. [45]	Tier 3	Men	15	24.4 ± 4.2	HRave (bpm), HRmax (%), and RPE (A.U.)	TD (m), TD at different speed thresholds, and acceleration and deceleration distance at > 2.5 m/s^2	n.a.
Asín-Izquierdo et al. [46]	Tier 2	Men	30	21.6 ± 4.4	HRave (bpm), HRmax (%), RPE (A.U.), and TRIMP (A.U.)	TD (m), TD at different speed thresholds, and acceleration and deceleration	n.a.
Beato et al. [47]	Tier 2	Men	12	22.5 ± 1.8	HRmax (%) and RPE (A.U.)	n.a.	Passes, dribbles, and passes in the target
Beato et al. [48]	Tier 3	Men	25	27 ± 9	RPE (A.U.)	TD (m/min), TD at different speed thresholds and acceleration and deceleration	n.a.
Beenham et al. [49]	Tier 2	n.d.	40	17.0 ± 0.6	n.a.	Player load	n.a.
Bergmann et al. [50]	Tier 2	Both	85	7 to 9	n.a.	n.a.	Total actions, passes, dribbles, goals, and duels
Bergmann et al. [51]	Tier 2	Both	43	9	n.a.	n.a.	Game performance, decision-making index, skill execution, and support index
Brandes et al. [52]	Tier 2	Men	17	14.9 ± 0.7	HR (bpm), HRmax (%), HR zones, RPE (A.U.), and BLa (mmol/L)	TD at different speed thresholds	n.a.

▶ **Table 3** Continued.

Study	Competitive level	Sex	*N*	Age (y)	Physiological measures extracted	Physical demands extracted	Technical actions extracted
Bravo-Sánchez et al. [53]	Tier 2	Men	154	10.7 ± 0.8	n.a.	TD, TD at different speed thresholds, and acceleration	Passes, dribbles, recoveries, clearances, shots, ball contacts, and attacking movements
Brito et al. [54]	Tier 2	n.d.	197	10.1 ± 2.7	n.a.	TD and TD at different speed thresholds	n.a.
Casamichana et al. [55]	Tier 2	n.d.	18	23.4 ± 4.5	n.a.	TD, TD at different speed thresholds, player load, maximum speed, work-to-rest ratio, and accelerations	n.a.
Castellano et al. [56]	Tier 3	Men	14	21.3 ± 2.3	n.a.	TD (m), player load, and work to rest ratio	n.a.
Castellano et al. [57]	Tier 2	n.d.	22	12.1 ± 0.4	HRmax (%) and HR zones	TD, player load, maximum speed, and work-to-rest ratio	n.a.
Clemente [58]	Tier 2	Men	24	16.3 ± 0.8	HRave (bpm) and RPE	TD (m/min)	Successful passes, unsuccessful passes, successful shots, unsuccessful shots, lost ball, and interceptions
Clemente et al. [59]	Tier 2	Men	12	7.6 ± 0.5	n.a.	n.a.	Received balls, conquered balls, lost balls, attacking balls, neutral balls, not-succeeded shots. and succeeded shots
Clemente et al. [60]	Tier 4	Men	22	24.6 ± 2.8	n.a.	TD (m/min), TD at different speed thresholds, and player load	n.a.
Clemente et al. [61]	Tier 2	Men	16	10.1 ± 0.3	n.a.	n.a.	Received balls, conquered balls, lost balls, attacking balls, neutral balls, not-succeeded shots, and succeeded shots
Clemente et al. [62]	Tier 4	Men	23	24.6 ± 2.8	n.a.	TD (m/min), TD at different speed thresholds, and accelerations	n.a.
Clemente et al. [63]	Tier 3	Men	23	16.8 ± 0.4	HR (bpm), HRave (bpm), and HRpeak (bpm)	TD (m/min), TD at different speed thresholds, and accelerations and decelerations	n.a.
Costa et al. [64]	Tier 2	Men	11	11.1 ± 0.4	n.a.	n.a.	Goals, shots, passes, player's involvement, ball possession, volume of play, and player's efficiency index and performance score
Da Silva et al. [65]	Tier 2	Men	16	13.5 ± 0.7	HRmax (%)	n.a.	Passes, target passes, involvements with the ball, crosses, dribbling, shot on goals, tackles, and headers
Dalen et al. [66]	Tier 3	Men	26	24.9 ± 4.2	n.a.	TD (m/min), TD at different speed thresholds, accelerations and decelerations, and player's load	n.a.

▶ **Table 3** Continued.

Study	Competitive level	Sex	*N*	Age (y)	Physiological measures extracted	Physical demands extracted	Technical actions extracted
Dellal et al. [67]	Tier 4	n.d.	20	27.4 ± 1.5	n.a.	TD (m) and TD at different speed thresholds	Duels, successful passes, balls lost, and possessions
Dellal et al. [68]	Tier 4	Men	10	26.0 ± 2.9	HRres (%)	n.a.	n.a.
Dellal et al. [69]	Tier 4	n.d.	20	27.0 ± 2.0	HRmax (%), HRres (%), RPE (A.U.), and BLa (mmol/L)	TD (m) and TD at different speed thresholds	Duels, successful passes, balls lost, and possessions
Dellal et al. [70]	Tier 4	n.d.	20	27.4 ± 1.5	HRmax (%), HRres (%), RPE (A.U.), and BLa (mmol/L)	TD (m) and TD at different speed thresholds	Duels, successful passes, balls lost, and possessions
Dellal et al. [71]	Tier 3	n.d.	27	16.5 ± 0.5	HRres (%)	n.a.	n.a.
Dimitriadis et al. [72]	Tier 2	n.d.	16	14.8 ± 0.5	HRmax (%) and RPE (A.U.)	TD (m), TD at different speed thresholds, and accelerations and decelerations	n.a.
Edis et al. [73]	Tier 2	Men	16	17.2 ± 1.0	HR (bpm)	n.a.	Successful and unsuccessful tackles, successful and unsuccessful shooting, successful and unsuccessful dribbling, and successful and unsuccessful pass
Evangelos et al. [74]	Tier 2	Men	9	17.2 ± 0.5	HRave (bpm), HRpeak (bpm), and BLa (mmol/L)	n.a.	Turn, dribble, tackle, interception, receive, pass, header, and block
Farhani et al. [75]	Tier 3	Men	16	20.7 ± 0.7	HRpeak (%), BLa (mmol/L), and RPE (A.U.)	n.a.	Successful passes, successful tackles, successful duels, and ball lost
Febré et al. [76]	Tier 2	Men	10	9.3 ± 0.4	HRmax (bpm), HRmean (bpm) and HR zones 1–5	n.a.	Successful and unsuccessful passes, interceptions, tackles, and steal
García-Angulo et al. [77]	Tier 2	Men	40	11.7 ± 0.4	n.a.	n.a.	Steal, clearance, rebound, pass interception, set-piece kick, pass, pass after throw-in, pass after set-piece, and throw-in
Giménez et al. [78]	Tier 3	Men	14	23.2 ± 2.7	n.a.	TD (m), TD at different speed thresholds and accelerations and decelerations at different thresholds	n.a.
Halouani et al. [79]	Tier 2	n.d	18	13.5 ± 0.7	HR (bpm), BLa (mmol/L), and RPE (A.U.)	n.a.	n.a.
Harrison et al. [80]	Tier 2	Men	10	13.0 ± 0.3	HRpeak (%), HR zones, and RPE (A.U.)	TD (m) and TD at different speed thresholds	Possessions, passes, successful passes, and shots: successful shots
Hill-Haas et al. [81]	Tier 2	Men	16	16.3 ± 0.6	HRmax (%), HR zones, BLa (mmol/L), and RPE (A.U.)	TD (m) and TD at different speed thresholds	n.a.
Jastrzębski et al. [82]	Tier 3	n.d	13	27.1 ± 5.2	HRmax (%)	TD (m)	n.a.
Köklü et al. [83]	Tier 3	Men	15	17.0 ± 1.0	HRmax (%), BLa (mmol/L), and RPE (A.U.)	TD (m), TD at different speed thresholds, and maximum speed	n.a.

▶ **Table 3** Continued.

Study	Competitive level	Sex	*N*	Age (y)	Physiological measures extracted	Physical demands extracted	Technical actions extracted
Köklü [84]	Tier 2	n.d	20	16.6 ± 0.5	HR (bpm), HRmax (%), and BLa (mmol/L)	n.a.	n.a.
Köklü et al. [85]	Tier 2	n.d	16	14.2 ± 0.6	HR (bpm), HRmax (%), and RPE (A.U.)	n.a.	n.a.
Köklü et al. [86]	Tier 2	n.d	14	16.7 ± 0.6	HRmax (%), RPE (A.U.), and BLa (mmol/L)	n.a.	n.a.
Köklü et al. [87]	Tier 2	n.d	16	15.7 ± 0.4	HR (bpm), HRmax (%), and BLa (mmol/L)	n.a.	n.a.
Köklü et al. [88]	Tier 2	n.d	16	16.5 ± 1.5	HRmax (%), RPE (A.U.), and BLa (mmol/L)	TD (m) and TD at different speed thresholds	n.a.
Köklü et al. [89]	Tier 2	Women	16	16.1 ± 2.3	HRmax (%), RPE (A.U.), and BLa (mmol/L)	TD (m) and TD at different speed thresholds	n.a.
Little et al. [90]	Tier 3	n.d	28	24.0 ± 5.0	HRmax (%) and RPE (A.U.)	n.a.	n.a.
Lex et al. [91]	Tier 2	n.d	n.d	7	n.a.	n.a.	Passing, dribbling, shooting, and scoring
López-Fernández et al. [92]	Tier 2	n.d	63	14–18	n.a.	TD (m), TD at different speed thresholds, maximum speed, and mean speed	n.a.
Makar et al. [93]	Tier 2	Men	20	17	n.a.	Peak speed and sprint	n.a.
Mara et al. [94]	Tier 3	Women	18	24.3 ± 4.2	HRpeak (%), HRmean (%), and HR zones	TD (m/min), TD at different speed thresholds, peak speed, and acceleration	n.a.
Martin-Garcia et al. [95]	Tier 3	n.d	21	20.4 ± 1.2	n.a.	TD (m/min), TD at different speed thresholds, and accelerations and decelerations	n.a.
Owen et al. [96]	Tier 4	Men	15	26.3 ± 4.9	HRmax (%) and HR zones	n.a.	Block, dribble, header, interception, pass, receive, shot, turn, tackle, and ball contacts
Owen et al. [13]	Tier 4	Men	10	27.6 ± 4.1	n.a.	TD (m) and TD at different speed thresholds	Block, dribble, header, interception, pass, receive, shot, turn, and tackle
Papadopoulos et al. [97]	Tier 3	Men	16	18.4 ± 2.3	HRmax (%)	TD (m), TD at different speed thresholds, and accelerations	n.a.
Pinheiro et al. [98]	Tier 5	Men	9	25.1 ± 4.6	n.a.	TD (m) and accelerations	n.a.
Rampinini et al. [99]	Tier 2	n.d	20	24.5 ± 4.1	HRmax (%), BLa (mmol/L), and RPE (A.U.)	n.a.	n.a.
Rebelo et al. [100]	Tier 2	Men	18	20.7 ± 1.0	HR (bpm) and BLa (mmol/L)	TD (m/min), TD at different speed thresholds and accelerations and decelerations	Tackles, ball interceptions, duels, passes, runs with ball, shots, and goals
Savolainen et al. [101]	Tier 3	Women	37	21.2 ± 2.6	HRmax (%), RPE (A.U.), and BLa (mmol/L)	TD (m/min), TD at different speed thresholds, and accelerations and decelerations	n.a.

▶ **Table 3** Continued.

Study	Competitive level	Sex	*N*	Age (y)	Physiological measures extracted	Physical demands extracted	Technical actions extracted
Silva et al. [102]	Tier 2	Men	20	16.8 ± 0.4	HRmean (bpm) and HRpeak (bpm)	TD (m/min), TD at different speed thresholds, and maximum speed	n.a.
Wang et al. [103]	Tier 2	Men	24	16.6 ± 0.5	HRmean (bpm) and RPE (A.U.)	n.a.	Successful and unsuccessful passes and lost ball
Yilmaz et al. [104]	Tier 2	Men	24	21.0 ± 1.5	RPE (A.U.)	n.a.	Successful and unsuccessful passes, lost ball, and interception
Younesi et al. [105]	Tier 3	Men	20	28.1 ± 4.6	HRave (bpm), HR zones, TRIMP	n.a.	n.a.
Younesi et al. [106]	Tier 3	Men	20	28.1 ± 4.6	n.a.	TD (m/min), TD at different speed thresholds, and mechanical work	n.a.

Abbreviations: BLa, blood lactate; F, female; HR, heart rate; HRave, average heart rate; HRpeak, peak heart rate; M, male; n.a., not applicable.; n.d., not described; RPE, rating of perceived exertion; T2, trained/developmental level (tier 2); T3, highly trained/national level (tier 3); T4, elite/international level (tier 4); T5, world class (tier 5); TRIMP, training impulse; TD: total distance.

observed for ratios of > 100 m^2 (Hedge's $g = -0.210$, $p = 0.88$), as shown in ▶ **Fig. 3**.

Finally, the sub-group analysis comparing 3v3 vs. 4v4 in terms of the rate of perceived exertion revealed that the 3v3 group was significantly favored over the 4v4 group for player-to-area ratios of > 100 m^2 (Hedge's $g = -0.820$, $p < 0.05$). However, for player-to-area ratios of ≤ 100 m^2, the ES favored the 3v3 group but lacked statistical significance (Hedge's $g = -0.178$, $p = 0.19$). The effect of the maximal HR was non-significant under both conditions. The total distance covered was significantly greater in the 4v4 format compared to the 2v2 format (Hedge's $g = 1.230$, $p < 0.05$). This ES was also influenced by the player-to-area ratio: for ratios of ≤ 100 m^2, differences between game formats were large (Hedge's $g = 1.714$, $p < 0.05$), whereas for ratios of > 100 m^2, no significant differences were noted (Hedge's $g = 0.265$, $p = 0.71$), as illustrated in ▶ **Fig. 4**.

Certainty of evidence

▶ **Table 8** presents the certainty of evidence in the GRADE evaluation, focusing on specific outcomes (RPE, BLa and ball possession) that exhibit more consistent result trends and similar comparisons (between 2v2 and 4v4) across multiple meta-analyses. Specific data on the certainty of evidence for multiple comparisons and physiological, physical, and technical outcomes can be found in the public repository of the review at https://osf.io/qza3h/. A very low certainty of evidence was observed across all meta-analyses conducted, primarily due to inconsistency, imprecision, and small sample sizes.

Discussion

Our systematic review with meta-analysis aimed to summarize the effects of playing formats on the physiological, physical, and technical demands commonly measured during sided games. Given the availability of multiple comparisons across different sided game formats, the meta-analysis conducted suggests that 2v2 is significantly more intense in terms of RPE and BLa concentrations. Similar evidence is observed for 3v3, which is more intense than the other sided game formats (4v4, 5v5, and 6v6). Although the maximal HR (HR max) and average did not show clear trends, significant differences were observed between 3v3 and 4v4 and 5v5, with 3v3 showing a greater intensity in both the HR max and the HR average.

Conversely, when examining physical demands, 4v4 and 3v3 formats tended to cover significantly greater distances across various speed thresholds compared to 2v2. However, beyond 4v4, the differences between formats became less pronounced.

Regarding technical outcomes, ball possession tended to be significantly higher in smaller formats, with 2v2 and 3v3 being more favorable for this outcome than 4v4. However, no significant effects were observed for lost balls or tackles. Interestingly, 3v3 had the highest number of successful shots and dribbles compared to larger formats such as 4v4, 5v5, and 6v6. For completed passes, no significant differences were found in most comparisons, except that 3v3 had a higher number than 2v2.

Physiological outcomes

Comparisons among 2v2, 3v3, and 4v4 were conducted through the meta-analysis, revealing that 2v2 was significantly more intense, producing higher BLa than the other formats, while 3v3 also showed higher values of this variable than 4v4. The results confirm that the smaller the format, the greater the anaerobic contribution to the activity, which can be inferred from the accumulated lactate levels. This may be a consequence of the fact that, generally, the smaller the format, the shorter the duration, which can particularly affect blood lactate responses, as it is measured at the end of exercise. It has been found that blood lactate levels can be influenced by anaerobic taxes, as the HR tends to increase linearly, while lac-

▶ **Table 4** Characteristics of the sided games included and compared.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Aasgaard et al. [36]	2v2	20×20 m	100	n.a.	Small goals	4	4	2
	3v3	30×21 m	105					
	4v4	33.3×26 m	108.1					
Abrantes et al. [37]	3v3+GK	30×20 m	100	Only defending or attacking	Goalkeepers	4	4	2
	4v4+GK	40×20 m	100	Regular game				
Ade et al. [38]	1v1	27×18 m	243	n.a.	Small goals	8	30 s	120 s
	2v2	27×18 m					60 s	60 s
Aguiar et al. [39]	2v2	20×21 m	150	n.a.	Small goals	3	6	1
	3v3	n.d.	150					
	4v4	40×30 m	150					
	5v5	n.d.	150					
Alashti et al. [40]	2v2	25×20 m	125	n.a.	Stop-ball rules	8	2	1
	4v4	35×28 m	122.5			4	4	2
Alcântara et al. [41]	3v3+GK	37×24 m	148	Influence of the regimen	Goalkeepers	4 or 2 or 1	5 or 10 or 20	150 to 300 s
	5v5+GK	48×31 m	148.8					
Almeida et al. [42]	3v3+GK	46×31 m	238209	n.a.	Goalkeepers	2	5	1
	6v6+GK	62×40.4 m						
Arslan et al. [43]	2v2	24×12 m	72	Passive vs. active rest	Ball possession	4	2	3
	3v3	30×18 m	90			4	3	3
	4v4	36×24 m	108			4	4	3
Aşçı et al. [44]	3v3+GK	25×20 m	83	n.a.	Goalkeepers	6	2	n.d.
	4v4+GK	35×30 m	131			4	3	
	5v5+GK	45×30 m	135			3	4	
	7v7+GK	55×40 m	157			2	6	
	9v9+GK	70×40 m	155			2	8	
Asian-Clemente et al. [45]	3v3	30×20 m	100	The use of floaters	Ball possession	4	4	2
	5v5	37×27 m	99.9					
	7v7	44×32 m	100.6					
	3v3+2	30×20 m	100					
	5v5+2	37×27 m	99.9					
	7v7+2	44×32 m	100.6					
Asín-Izquierdo et al. [46]	3v3 and 3v3+GK	20/30/40×20 m	67/90/160	Feedback vs no feedback, goalkeepers vs no, training regimen, and field	Goalkeepers	1 or 3	6 or 2	1
	6v6 and 3v3+GK	40×30 or 30×40 m	100		n.d.			

▶ **Table 4** Continued.

▸ **Table 4** Characteristics of the sided games included and compared.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Beato et al. [47]	3v3 + GK	30 × 20 m	100	n.a.	Goalkeepers	6	4	2
	4v4 + GK	30 × 20 m	75					
Beato et al. [48]	2v2 + GK		121	n.a.	Goalkeepers			
	3v3 + GK		72					
	4v4 + GK		68 and 104					
	5v5 + GK		72 and 115					
	6v6 + GK		67 and 140					
	7v7 + GK		157.5					
	8v8 + GK		229 and 353.4					
	10v10 + GK							
Beenham et al. [49]	2v2	20 × 15 m	300	n.a.	Ball possession	4	2	3
	3v3	25 × 18 m	450			4	3	3
	4v4	30 × 20 m	600			4	4	3
Bergmann et al. [50]	5v5 + GK	28 × 20 m	56	Training regimen	Goalkeepers	4	5	n.d.
	7v7 + GK	45 × 35 m	125.5			1	20	
Bergmann et al. [51]	5v5 + GK	28 × 20 m	56	Training regimen	Goalkeepers	2	20	n.d.
	7v7 + GK	45 × 35 m	125.5					
Brandes et al. [52]	2v2	28 × 21 m	147	n.a.	Small goals	3	4	1.5
	3v3	34 × 26 m	147			3	5	
	4v4	40 × 30 m	150			3	6	
Bravo-Sánchez et al. [53]	7v7 + GK	60 × 40 m	171	n.a.	Goalkeepers	2	25	5
	8v8 + GK	60 × 40 m	150					
Brito et al. [54]	5v5 + GK	45.5 × 29 m	132	n.a.	Goalkeepers	1	30	n.a.
	7v7 + GK	64 × 41 m	187					
	9v9 + GK	82 × 52 m	237					
Casamichana et al. [55]	3v3	29 × 19 m	92	Man vs. non-man marking	Ball possession	1	6	n.a.
	6v6	40 × 28 m	93					
	9v9	55 × 30 m	92					
Castellano et al. [56]	3v3 and 3v3 + GK	43 × 30 m	215	Scoring method	Goalkeepers	1	6	n.a.
	5v5 and 5v5 + GK	55 × 38 m	209		Small goals			
	7v7 and 7v7 + GK	64 × 46 m	210		Ball possession			
Castellano et al. [57]	7v7 + GK	n.d.	100, 200 and 300	n.a.	Goalkeepers	2	12	5
	9v9 + GK	n.d.	100, 200 and 300					

▶ **Table 4** Continued.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Clemente [58]	2v2	30×15 m	112.5	n.a.	Small goals	4	2	2
	4v4	35×25 m	109.4			4	4	2
Clemente et al. [59]	3v3	20×15 m	50	n.a.	Small goals	3	3	2
	6v6	30×22 m	55					
Clemente et al. [60]	5v5+2	40×31 m	103	Scoring method	Goalkeepers	2	6	3
	5v5+GK	40×31 m	124		Ball possession	2	6	3
	10v10+2	52×44 m	104			2	10	3
	10v10+GK	52×44 m	114			2	10	3
Clemente et al. [61]	3v3	20×15 m	50	n.a.	Small goals	3	3	2
	6v6	30×22 m	55			3	6	1
Clemente et al. [62]	5v5+GK	40×31 m	124	n.a.	Goalkeepers	2	6	3
	6v6+GK	45×32 m	120			3	7.5	3
	9v9+GK	70×50 m	194			2	11	3
Clemente et al. [63]	3v3	39×24 and 32×19 m	156 and 101	n.a.	Small goals	1	3	n.a.
	5v5	50×31 and 40×25 m	155 and 100			1	5	n.a.
Costa et al. [64]	5v5+GK	66×43 m	284	n.a.	Goalkeepers	2	25	n.d.
	7v7+GK	82×52 m	305					
Da Silva et al. [65]	3v3	30×30 m	150	n.a.	Small goals	3	4	3
	4v4	30×30 m	113					
	5v5	30×30 m	90					
Dalen et al. [66]	4v4+GK	39×39 m	190	n.a.	Goalkeepers	6–8	3	3
	6v6+GK	47×43 m	168			3–5	5	2
Dellal et al. [67]	2v2+4	20×15 m	75	Touches limitation	Ball possession	4	2	n.d.
	3v3+4	25×18 m	75			4	3	
	4v4+4	30×20 m	75			4	4	
Dellal et al. [68]	1v1	10×10 m	50	Scoring method	Goalkeepers	4	1.5	1.5
	2v2	20×20 m	100		Ball possession	6	2.5	1.5
	4v4+GK	30×25 m	94			2	4	3
	8v8+GK	60×45 m	169			2	10	5
	8v8	60×45 m	169			4	4	3
	10v10+GK	90×45 m	203			3	20	5
Dellal et al. [69]	2v2+4	20×15 m	75	n.a.	Ball possession	4	2	3
	3v3+4	25×18 m	75			4	3	3
	4v4+4	30×20 m	75			4	4	3

▶ **Table 4** Continued.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Dellal et al. [70]	2v2	20×15 m	75	n.a.	Ball possession	4	2	3
	3v3	25×18 m	75			4	3	3
	4v4	30×20 m	75			4	4	3
Dellal et al. [71]	2v2	25×20 m	125	n.a.	Stop-ball rules	8	2	1
	3v3	30×25 m	125			6	3	1.5
	4v4	35×28 m	122.5			4	4	2
Dimitriadis et al. [72]	1v1 + GK	15×10, 20×10 and 20×15 m	75, 100 and 150	Field dimensions	Goalkeepers	4	1	2
	2v2 + GK	20×15, 27×15 and 30×20 m	75, 100 and 150			4	2	4
	3v3 + GK	25×18, 30×20 and 36×25 m	75, 100 and 150			4	3	3
Edis et al. [73]	1v1	15×10 m	38	n.a.	n.d.	4	1	1
	2v2	24×18 m	72			4	2	2
	3v3	40×30 m	150			4	3	3
Evangelos et al. [74]	3v3	25×20 m	83	n.a.	Ball possession	1	3	12
	3v3 + 1DF	25×20 m	83					
	3v3 + 1AT	25×20 m	83					
	4v3	25×20 m	71					
	4v4	30×25 m	94					
	4v4 + 1DF	30×25 m	94					
	4v4 + 1AT	30×25 m	94					
	5v4	30×25 m	83					
Farhani et al. [75]	3v3	30×20 m	100	Training regimen	Small goals	1/2/3	12/6/4	n.d.
	4v4	32×25 m	100					
Febré et al. [76]	3v3	30×30 m	150	n.a.	n.d.	3	4	3
	4v4	30×30 m	112.5					
	5v5	30×30 m	90					
García-Angulo et al. [77]	5v5 + GK	38×20 m	190	n.a.	Goalkeepers	2	20	n.d.
	8v8 + GK	58×38 m	314					
Giménez et al. [78]	4v4 + GK	30×30 m	113	n.a.	Goalkeepers	8	4	2
	8v8 + GK	42.4×42.4 m	113					
Halouani et al. [79]	2v2	25×20 m	125	Scoring method	Stop-ball rules	4	4	2
	3v3	25×20 m	83		Small goals			
	4v4	25×20 m	63					
Harrison et al. [80]	3v3	35×25 m	146	n.a.	Small goals	1	16	n.a.
	6v6	49×35 m	143					

▶ **Table 4** Continued.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Hill-Haas et al. [81]	2v2	28 × 21 m	150	n.a.	Small goals	1	24	n.a.
	4v4	40 × 30 m	150					
	6v6	49 × 37 m	150					
Jastrzębski et al. [82]	4v4 + GK	40 × 30 m	150	n.a.	Goalkeepers	4	4	2
	5v5 + GK	43 × 33 m	142					
Köklü et al. [83]	2v2	25 × 16 m	100	Training regimen	Ball possession	1/6/3/2	12/2/4/6	n.a./2/2/2
	3v3	30 × 20 m	100					
	4v4	32 × 25 m	100					
Köklü [84]	2v2	20 × 15 m	75	Training regimen	n.d.	3/0	2/6	n.d.
	3v3	24 × 18 m	72			3/0	3/9	
	4v4	36 × 24 m	108			3/0	4/12	
Köklü et al. [85]	3v3 + 4	20 × 15, 25 × 18 and 30 × 20 m	50, 75 and 100	Field dimensions	Ball possession	4	3	2
	4v4 + 4	20 × 20, 30 × 20 and 32 × 25 m	50, 75 and 100			4	4	2
Köklü et al. [86]	2v2	24 × 12 m	72	n.a.	Ball possession	8	2	2
	3v3	30 × 18 m	90			8	3	2
	4v4	36 × 24 m	108			8	4	2
Köklü et al. [87]	1v1	18 × 6 m	54	n.a.	n.d.	6	1	2
	2v2	24 × 12 m	72			6	2	2
	3v3	30 × 18 m	90			6	3	2
	4v4	36 × 24 m	108			6	4	2
Köklü et al. [88]	2v2	27 × 15 m	100	n.a.	Ball possession	4	2	2
	3v3	30 × 20 m	100			4	3	2
	4v4	32 × 25 m	100			4	4	2
Köklü et al. [89]	2v2 + 4	20 × 20 m	100	n.a.	n.d.	4	2	2
	3v3 + 4	30 × 20 m	100			4	3	2
	4v4 + 4	40 × 20 m	100			4	4	2
Little et al. [90]	2v2	30 × 20 m	150	n.a.	n.d.	4	2	2
	3v3	43 × 25 m	179			4	3.5	1.5
	4v4	40 × 30 m	150			4	4	2
	5v5	45 × 30 m	135			4	6	1.5
	6v6	50 × 30 m	125			3	8	1.5
	8v8	70 × 45 m	197			4	8	1.5
Lex et al. [91]	3v3 + GK	n.d.	n.d.	n.a.	Goalkeepers	n.d.	10 and 12	n.d.
	6v6 + GK	55 × 35 m	160			2	20	10
	7v7 + GK	55 × 35 m	138					

▶ **Table 4** Continued.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
López-Fernández et al. [92]	3v3	30×20 m	100	n.a.	n.d.	1	4	4
	4v4	32×25 m	100					
	5v5	37×27 m	99.9					
	6v6	40×30 m	100					
Makar et al. [93]	1v1	15×10 m	75	n.a.	Small goals	4	0.5	2
	5v5	40×30 m	120			4	4	2
Mara et al. [94]	4v4+GK	40×40 m	200	n.a.	Goalkeepers	2	5	2
	5v5+GK	50×40 m	200			2	5	2
	6v6+GK	60×40 m	200			2	7	2
	7v7+GK	70×40 m	200			2	7	2
	8v8+GK	80×68 m	340			2	9	2
	9v9+GK	90×68 m	340			2	9	2
Martin-Garcia et al. [95]	5v5+GK	40×33 m	132	n.a.	Goalkeepers	n.d.	5.5	n.d.
	9v9+GK	72×65 m	260			n.d.	12	n.d.
Owen et al. [96]	3v3+GK	30×25 m	125	n.a.	Goalkeepers	3	5	4
	9v9+GK	60×50 m	166.6					
Owen et al. [13]	4v4+GK	30×25 m	94	n.a.	Goalkeepers	3	5	n.d.
	5v5+GK	46×40 m	184					
	6v6+GK	50×44 m	183					
	7v7+GK	54×45 m	174					
	8v8+GK	60×50 m	188					
	9v9+GK	70×56 m	218					
	10v10+GK	80×70 m	280					
Papadopoulos et al. [97]	1v1	10×15, 20×10 and 20×15 m	75, 100 and 150	Field dimensions	Small goals	4	1	1
	2v2	20×15, 27×15 and 30×20 m	75, 100 and 150			4	2	2
	3v3	25×18, 30×20 and 36×25 m	75, 100 and 150			4	3	3
	4v4	30×20, 35×25 and 40×30 m	75, 100 and 150			4	4	4
Pinheiro et al. [98]	4v4+GK	32×20 m	80	n.a.	Goalkeepers	4	7.5	5
	6v6+GK	40×30 m	100			4	12.5	5
	7v7+GK	68×52.5 m	255			4	10	5
	8v8+GK	40×35 m	87.5			4	7.5	5
	10v10+GK	105×68 m	357			4	7.5	5

▶ **Table 4** Continued.

Study	Formats	Field dimensions	Player to area ratio (m^2)	Concurrent task constraints	Task objective	Repetitions[a]	Duration (min)[a]	Rest/recovery (min)
Rampinini et al. [99]	3v3	20×12, 25×15 and 30×18 m	40,63 and 90	Field dimensions	Small goals	3	4	3
	4v4	24×16, 30×20 and 36×24 m	48, 75 and 108	Couch encouragement				
	5v5	28×20, 35×25 and 42×30 m	56, 88 and 126					
	6v6	32×24, 40×30 and 48×36 m	64, 100 and 144					
Rebelo et al. [100]	4v4+GK	47.7×29.5 m	176	n.a.	Goalkeepers	6	6	3–5
	8v8+GK	85.9×53.2 m	286			2	18	3
Savolainen et al. [101]	4v4+GK	32×22 m	88	n.a.	Goalkeepers	5	3	3
	8v8+GK	75×48 m	225			3	5	3
Silva et al. [102]	3v3	39×24 and 32×19 m	156 and 101	n.a.	Small goals	1	3	n.a.
	5v5	50×31 and 40×25 m	155 and 100			1	5	n.a.
Wang et al. [103]	2v2	25×15 m	94	n.a.	Small goals	4	2	2
	4v4	35×17 m	92			4	4	2
Yilmaz et al. [104]	2v2	27×15 m	101	Scoring method	Goalkeeper	4	2	2
	2v2+GK	27×15 m	101		Small goals	4	4	2
	4v4	32×25 m	100		Ball possession			
	4v4+GK	32×25 m	100					
Younesi et al. [105]+	3v3	30×27 m	135	Scoring method and touches limitation	Goalkeeper	2	3	2
	3v3+GK	30×27 m	135		Ball possession	2	4	2
	4v4	32×22 m	88			2	6	2
	4v4+GK	32×22 m	88					
	6v6	40×28 m	59					
	6v6+GK	40×28 m	59					
Younesi et al. [106]	3v3	30×27 m	135	Scoring method and touches limitation	Goalkeeper Ball possession	2	3	2
	3v3+GK	30×27 m	135			2	4	2
	4v4	32×22 m	88			2	6	2
	4v4+GK	32×22 m	88					
	6v6	40×28 m	59					
	6v6+GK	40×28 m	59					

Abbreviations: AT, attacker; DF, defender; GK, goalkeepers; n.a., not applicable; n.d.m not described.

[a]Per session.

▶ **Table 5** Summary of the meta-analysis parameters of different formats of SSGs in relation to physiological outcomes, including the rate of perceived exertion, maximal HR, average HR, and blood lactate levels.

Outcome	SSG comparison			Meta-analysis					Hetero-geneity	Egger regression intercept		DT trim and fill			
		N studies	*N* partici-pants	Hedges's *g*	Lower limit	Upper limit	*p*-value	Favours	I^2 (%)	Intercept	*p*-Value	Trimmed studies	Adjusted value	Lower limit	Upper limit
RPE	2v2 vs. 3v3	15	258	−0.836	−1.208	−0.484	<0.01	2v2	74.7	−6.215	<0.05	0			
	2v2 vs. 4v4	18	330	−1.417	−1.994	−0.950	<0.01	2v2	87.7	−9.259	<0.01	1	−1.367	−1.859	−0.880
	2v2 vs. 5v5	3	63	−1.256	−1.633	−0.879	<0.01	2v2	0						
	3v3 vs. 4v4	17	281	−0.524	−0.819	−0.228	<0.01	3v3	66.4	−7.769	<0.01	0			
	3v3 vs. 5v5	6	109	−0.760	−1.462	−0.058	<0.05	3v3	84						
	3v3 vs. 6v6	5	113	−1.365	−1.872	−0.858	<0.01	3v3	66.1						
	4v4 vs. 5v5	4	83	−0.892	−1.206	−0.577	<0.01	4v4	0						
	4v4 vs. 6v6	3	73	−1.917	−2.541	−1.293	<0.01	4v4	60.3						
	5v5 vs. 6v6	3	73	−0.710	−1.041	−0.380	<0.01	5v5	0						
HRmax	1v1 vs. 2v2	4	60	0.459	−0.105	1.024	0.11	2v2	59.1						
	1v1 vs. 3v3	3	48	0.355	−0.809	1.519	0.55	3v3	87.5						
	1v1 vs. 4v4	3	48	0.341	−0.654	1.336	0.52	4v4	83.2						
	2v2 vs. 3v3	16	258	0.108	−0.345	0.561	0.64	3v3	84.5	−2.435	0.33	1	0.198	−0.269	0.667
	2v2 vs. 4v4	18	298	−0.267	−0.748	0.213	0.28	2v2	87.5	−6.944	0.05	0			
	3v3 vs. 4v4	23	414	−0.467	−0.739	−0.195	<0.01	3v3	74.3	−4.662	<0.01	0			
	3v3 vs. 5v5	9	202	−0.443	−0.931	0.046	0.07	3v3	82.1						
	3v3 vs. 6v6	5	139	−0.835	−1.825	0.155	0.10	3v3	92.5						
	3v3 vs. 7v7	3	51	−0.181	−1.049	0.688	0.68	3v3	83.4						
	4v4 vs. 5v5	7	183	−0.365	−0.603	−0.217	<0.01	4v4	20.3						
	4v4 vs. 6v6	3	111	−0.856	−1.558	−0.154	<0.05	4v4	81.8						
	5v5 vs. 6v6	3	111	−0.355	−0.668	−0.042	<0.05	5v5	20.8						
	5v5 vs. 7v7	3	51	−0.317	−0.901	0.267	0.29	5v5	55.2						
HRaverage	2v2 vs. 3v3	7	115	0.185	−0.398	0.768	0.535	3v3	79.5						
	2v2 vs. 4v4	9	159	−1.187	−2.124	−0.250	<0.05	2v2	92.6						
	3v3 vs. 4v4	14	271	−0.605	−1.012	−0.198	<0.01	3v3	79.1	−5.427	<0.01	0			
	3v3 vs. 5v5	7	164	−0.343	−0.733	0.047	0.081	3v3	64.4						
	3v3 vs. 6v6	4	131	−0.172	−1.016	0.672	0.689	3v3	89.9						
Blood lactate	2v2 vs. 3v3	13	224	−0.767	−1.005	−0529	<0.01	2v2	34.2	−9.988	<0.01	3	−0.766	−1.005	−0.528
	2v2 vs. 4v4	12	209	−1.599	−2.233	−0.965	<0.01	2v2	87.4	−12.729	<0.01	1	−1.741	−2.405	−1.078
	3v3 vs. 4v4	17	284	−0.852	−1.244	−0.460	<0.01	3v3	79.0	−14.856	<0.01	2	−0.979	−1.371	−0.586

Abbreviations: DT, Duval and Tweedie's trim and fill; HR, heart rate; ; RPE, rating of perceived exertion; SSG, small-sided games.

▶ **Table 6** Summary of the meta-analysis parameters of different formats of SSGs in relation to physical outcomes, including the total distance covered, maximal speed, distance covered (0–6.9 km/h), distance covered (7–13 km/h), distance covered (13–17.8 km/h), distance covered (> 18 km/h), distance covered (> 19 km/h), number of accelerations, number of decelerations and number of sprints.

Outcome	SSG comparison			Meta-analysis					Hetero-geneity	Egger regression intercept		DT trim and fill			
		N studies	*N* partici-pants	Hedg-es's *g*	Lower limit	Upper limit	*p*-value	Favours	I^2 (%)	Intercept	*p*-value	Trimmed studies	Adjust-ed value	Lower limit	Upper limit
Total distance covered	1v1 vs. 2v2	3	44	1.371	−0.608	3.350	0.18	2v2	93.7						
	2v2 vs. 3v3	12	198	2.635	0.666	1.329	<0.01	3v3	95.6	11.639	<0.01	0			
	2v2 vs. 4v4	13	214	4.095	2.585	5.604	<0.01	4v4	96.7	11.132	<0.01	0			
	3v3 vs. 4v4	14	282	1.299	0.581	2.018	<0.01	4v4	94.2	8.669	<0.01	0			
	3v3 vs. 5v5	7	167	0.468	−0.276	1.213	0.21	5v5	89.6						
	3v3 vs. 6v6	6	168	0.111	−0.797	1.098	0.81	6v6	93.2						
	4v4 vs. 5v5	5	121	−0.932	−2.090	0.226	0.12	4v4	92.0						
	4v4 vs. 6v6	6	161	0.141	−1.126	1.409	0.83	6v6	95.1						
	4v4 vs. 8v8	4	70	0.188	−2.372	2.748	0.89	8v8	96.0						
	5v5 vs. 6v6	4	117	−0.407	−1.010	0.196	0.19	5v5	77.1						
	5v5 vs. 7v7	4	98	0.523	−0.127	1.173	0.12	7v7	78.4						
	5v5 vs. 9v9	4	102	0.535	−0.495	1.565	0.31	9v9	91.3						
	5v5 vs. 10v10	4	78	1.207	0.532	1.882	<0.01	10v10	73.5						
	6v6 vs. 9v9	4	72	1.483	0.167	2.800	<0.05	9v9	91.1						
	6v6 vs. 10v10	3	45	0.743	0.091	1.395	<0.05	10v10	48.9						
	7v7 vs. 8v8	3	189	−0.253	−1.370	0.864	0.66	7v7	92.5						
Maximal speed	3v3 vs. 4v4	3	86	0.004	−0.301	0.309	0.98	4v4	2.2						
	3v3 vs. 5v5	4	117	0.456	0.050	0.862	<0.01	5v5	50.1						
Distance covered 0–6.9 km h^{-1}	2v2 vs. 3v3	8	111	1.838	0.723	2.952	<0.01	3v3	91.4						
	2v2 vs. 4v4	8	112	2.130	0.190	3.071	<0.01	4v4	96.4						
	3v3 vs. 4v4	9	176	1.289	0.529	2.049	<0.01	4v4	89.3						
	3v3 vs. 5v5	3	88	0.546	0.078	1.015	<0.05	5v5	41.1						
Distance covered 7–13 km h^{-1}	2v2 vs. 3v3	5	71	0.995	−0.469	2.460	0.18	3v3	93.2						
	2v2 vs. 4v4	13	214	2.525	0.845	4.205	<0.01	4v4	92.7	12.826	0.08	0			
	3v3 vs. 4v4	4	56	1.503	0.537	2.468	<0.01	4v4	80.3						
Distance covered 13–17.8 km h^{-1}	2v2 vs. 3v3	8	131	1.349	0.580	2.218	<0.01	3v3	85.6						
	2v2 vs. 4v4	5	72	2.049	1.087	3.123	<0.01	4v4	89.9						
	3v3 vs. 4v4	8	130	0.574	0.124	1.025	<0.05	4v4	67.9						
Distance covered > 18 km h^{-1}	2v2 vs 3v3	7	113	1.031	0.424	1.638	<0.01	3v3	76.7						
	2v2 vs 4v4	8	129	1.314	0.623	2.005	<0.01	4v4	82.5						
	3v3 vs 4v4	8	123	0.454	−0.046	0.954	0.08	4v4	73.6						

▲ **Table 6** Continued.

Outcome	SSG comparison			Meta-analysis					Hetero-geneity	Egger regression intercept		DT trim and fill			
		N studies	*N* partici-pants	Hedg-es's *g*	Lower limit	Upper limit	*p*-value	Favours	I^2 (%)	Intercept	*p*-value	Trimmed studies	Adjust-ed value	Lower limit	Upper limit
Distance covered > 19 km h^{-1}	3v3 vs. 4v4	3	101	0.310	−0.521	1.140	0.47	4v4	85.4						
	3v3 vs 5v5	3	103	−0.032	−0.573	0.509	0.91	3v3	68.7						
Number of accelerations	2v2 vs. 3v3	3	57	−0.047	−0.659	0.565	0.88	2v2	62.4						
	2v2 vs. 4v4	3	47	0.915	−0.200	2.029	0.11	4v4	86.3						
	3v3 vs. 4v4	4	120	0.549	−0.155	1.252	0.13	4v4	83.4						
	3v3 vs. 5v5	4	123	−0.135	−0.823	0.553	0.70	3v3	84.3						
Number of decelerations	3v3 vs. 4v4	3	95	0.489	−0.471	1.431	0.32	4v4	87.0						
	3v3 vs. 5v5	3	98	−0.178	−0.959	0.603	0.66	3v3	83.4						
Number of sprints	2v2 vs. 3v3	4	59	0.204	−0.386	0.795	0.50	3v3	62.1						
	2v2 vs. 4v4	5	75	0.418	0.102	0.735	< 0.05	4v4	0						
	3v3 vs. 4v4	5	115	0.060	−0.336	0.457	0.77	4v4	48.7						

Abbreviations: DT, Duval and Tweedie's trim and fill; SSG, small-sided games.

tate levels tend to escalate exponentially [99]. This may explain why, in our meta-analysis, the HR measures did not show significant differences between the smaller formats, while blood lactate levels did. The combination of short durations, increased individual participation, and limited rest periods creates conditions for an extremely intensified experience in games like 2v2 and 3v3, leading to higher anaerobic contribution in a short period of time. For example, in 2v2 games, accumulated levels often exceeded 8 $mmol \cdot L^{-1}$, as shown in previous studies [43, 86, 88, 89, 107]. However, certain mediators or moderators of intensification should be considered. For example, not using small goals and instead implementing a crossing line, where the point is achieved whenever a player crosses it, can influence the level of intensification. In such cases, the need for players to approach more frequently and create progression lines mediates intensification, as there are no moments of shooting or rest during transitions. Similarly, comparing the continuous play to stoppages can also impact the intensity of the game. In the continuous play, there is no pause for rest or regrouping, leading to higher intensity and a denser exercise load within each repetition. In contrast, games with stoppages allow brief recovery periods, which can reduce the overall intensity.

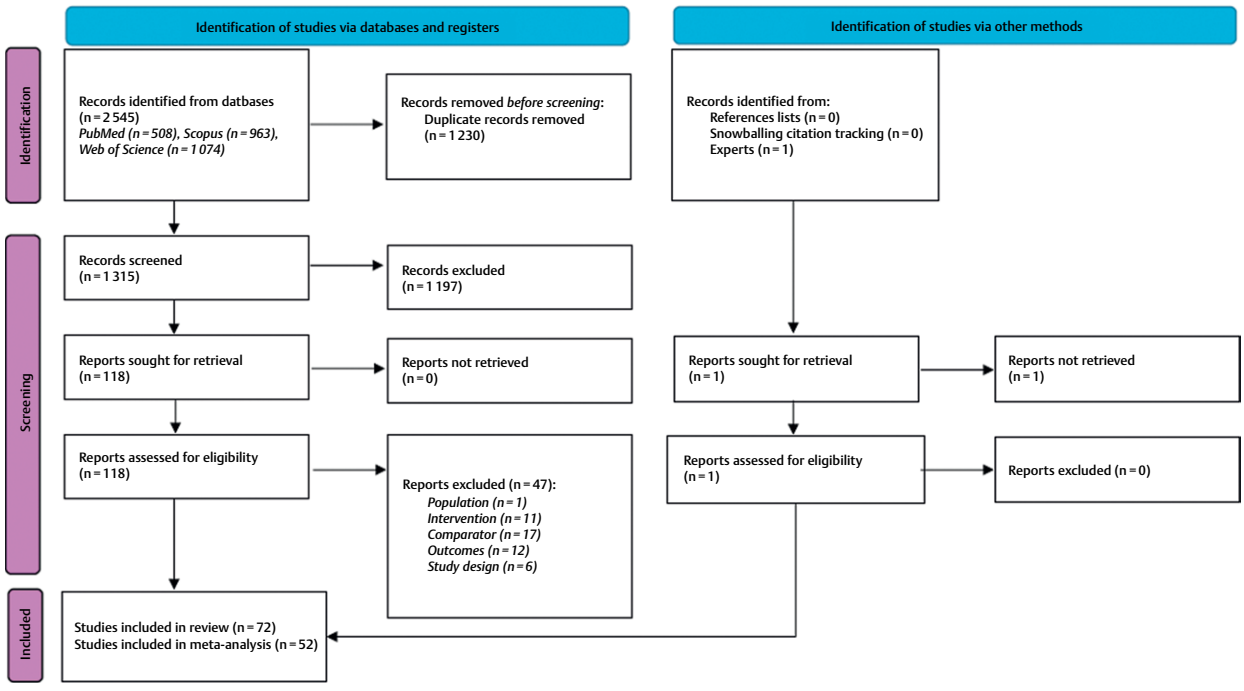
Following the trend observed in lactate levels, the meta-analysis of RPE also revealed higher intensity levels reported by players in smaller formats. Specifically, there was a clear progression in intensity: 2v2 was the most intense, followed by 3v3, which was more intense than the larger formats (i.e., 4v4 and 5v5 [69, 72, 81, 107]). Among the larger formats, 4v4 was more intense than 5v5 [48, 90, 99]. This pattern of increasing intensity from 5v5 to 4v4 to 3v3 and finally to 2v2 is consistent with the strong relationship between RPE and lactate levels. Previous research on sided games has shown a close correlation between RPE and lactate levels, although the most significant determinant for explaining RPE was the association between the HR and lactate levels [108].

Interestingly, among the physiological outcomes, the HR showed the least significant trends. This highlights concerns about the reliability of this indicator as a measure of load in acyclic tasks with frequent changes in activity. For example, no differences were observed in the HR max or average between any of the formats from 1v1 to 3v3 [72, 87, 97]. However, when comparing 3v3 to 4v4, significantly higher HR max and average values were observed in 3v3 [69, 71, 74, 89]. Furthermore, 4v4 was significantly more intense than 5v5 or 6v6 in terms of HR max [44, 90, 99]. Possibly, in smaller formats such as 1v1 to 3v3, the more rapid elevation in the HR may be due to an increased demand for oxygen and energy since the short duration of drills and the greater individuals' participation [58], leading to a greater contribution of the anaerobic energy systems, which contributes to higher lactate production [89]. This could help explain the findings discussed earlier in lactate levels. In larger formats like 4v4 and 5v5, HR kinetics may be slower because players have more space and slightly more recovery time between bursts of effort. Although the HR still increases, it does so at a less rapid pace compared to smaller formats, possibly resulting in less lactate accumulation during exercise.

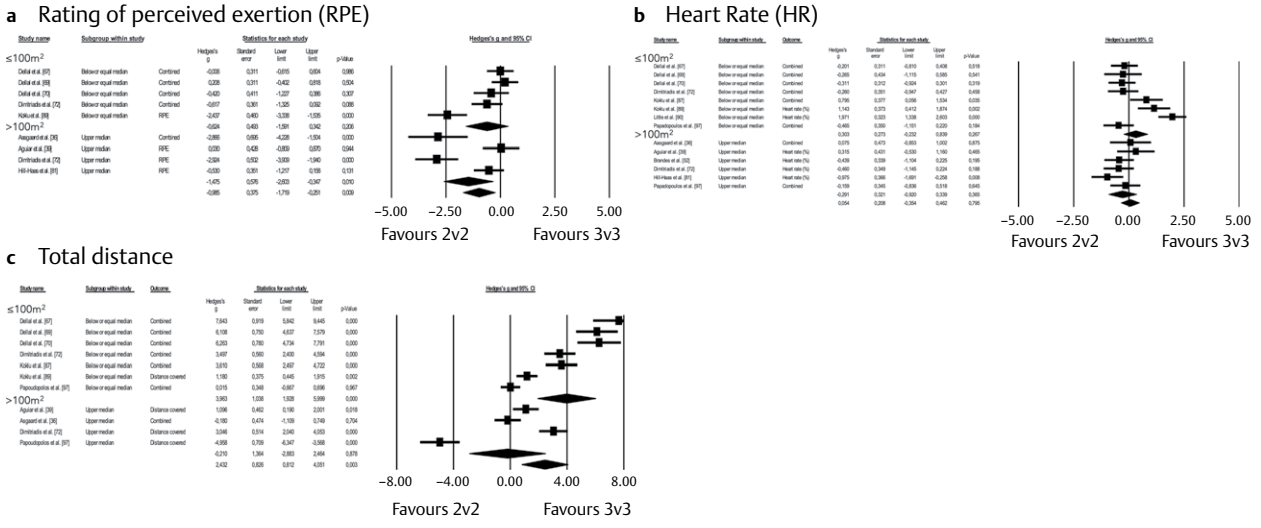
▶ **Table 7** Summary of the meta-analysis parameters of different formats of SSGs in relation to technical outcomes, including passes completed, unsuccessful passes, dribble, successful shots, tackles, ball possession and lost balls.

Outcome	SSG comparison			Meta-analysis					Heterogeneity	Egger regression intercept		DT trim and fill			
		N studies	*N* participants	Hedges's *g*	Lower limit	Upper limit	*p*-value	Favours	I^2 (%)	Intercept	*p*-value	Trimmed studies	Adjusted value	Lower limit	Upper limit
Passes completed	2v2 vs. 3v3	3	60	1.175	0.746	2.804	0.01	3v3	81.5						
	2v2 vs. 4v4	5	108	0.539	−0.710	1.789	0.40	4v4	94.2						
	3v3 vs. 4v4	8	144	0.051	−0.432	0.533	0.84	4v4	76.4						
	3v3 vs. 6v6	3	54	−0.432	−0.462	−1.338	0.35	3v3	80.3						
Unsuccessful passes	2v2 vs. 4v4	3	72	0.522	−1.683	2.728	0.64	4v4	96.3						
Dribble	3v3 vs. 4v4	3	41	−3.939	−7.290	−0.588	0.02	3v3	96.2						
Successful shoots	3v3 vs. 5v5	3	55	−0.763	−1.302	−0.225	<0.01	3v3	39.6						
	3v3 vs. 6v6	4	54	−1.060	−1.798	−0.322	<0.01	3v3	70.1						
Tackles	3v3 vs. 4v4	3	58	0.317	−0.083	0.717	0.12	4v4	0						
Ball possession	2v2 vs. 3v3	3	60	0.079	−0.414	0.571	0.75	3v3	46.9						
	2v2 vs. 4v4	3	60	−1.948	−2.690	−1.206	<0.01	2v2	63.9						
	3v3 vs. 4v4	3	60	−1.593	−2.005	−1.181	<0.01	3v3	0						
Lost balls	2v2 vs. 3v3	3	60	−1.466	−3.365	0.434	0.13	2v2	94.7						
	2v2 vs. 4v4	5	108	−1.411	−3.175	0.354	0.12	2v2	96.9						
	3v3 vs. 4v4	4	76	−0.230	−1.636	1.176	0.75	3v3	93.5						

Abbreviations: DT, Duval and Tweedie's trim and fill; SSG, small-sided games.



▶ **Fig. 1** PRISMA 2020 flow diagram [14]. PRISMA, preferred reporting items for systematic reviews and meta-analyses.

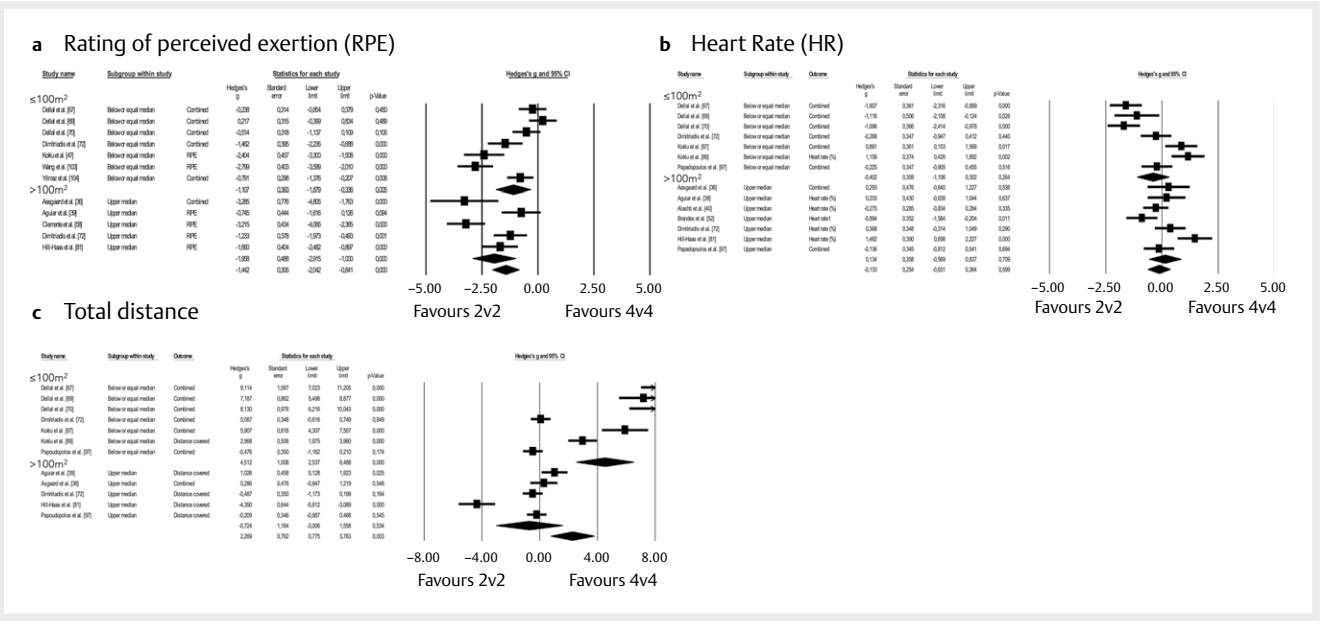


▶ **Fig. 2** Comparisons between 2v2 and 3v3 formats based on player-to-area ratios. CI, confidence interval.

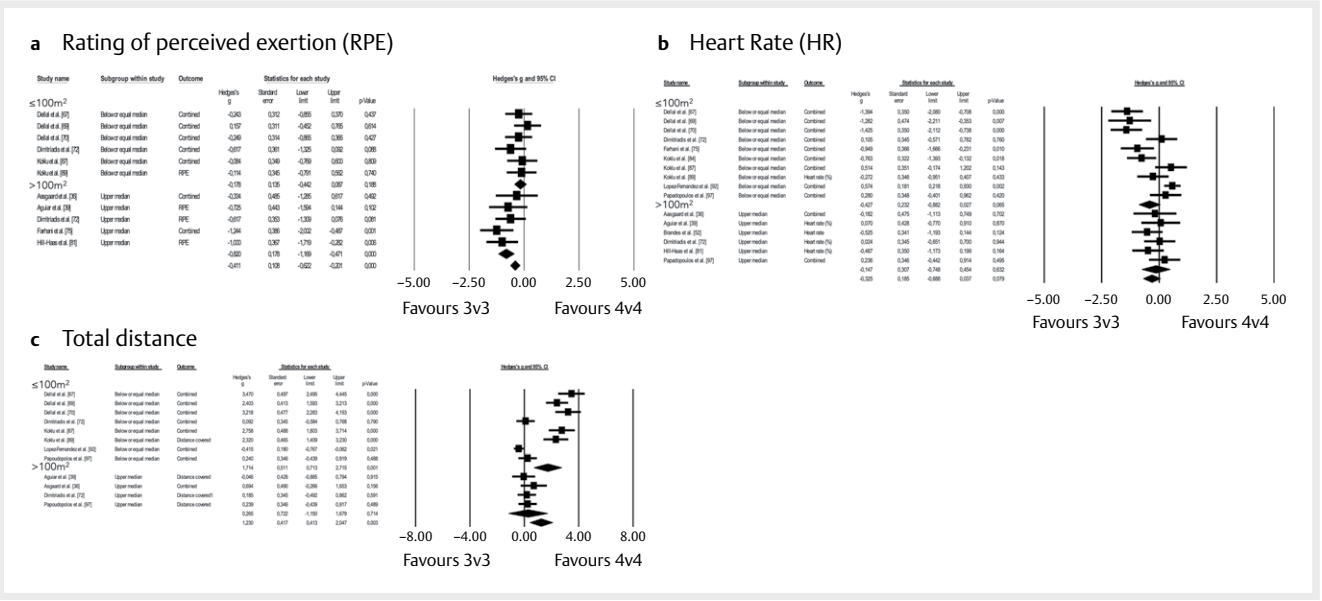
Physical outcomes

The total distances covered in different formats, ranging from 2v2 to 4v4, revealed that the 4v4 format required significantly greater distances to cover [39, 48, 71, 72, 81, 88, 89, 107], with the 3v3 format also involving significantly more running than the 2v2 format [36, 39, 48, 72]. These results suggest that, in smaller formats, the physiological effort may often be more related to the frequency of other actions rather than the total amount of running. In tightly confined spaces, such as those used in 2v2, there may be limited opportunities for running [109]. In contrast, larger spaces, like those in 4v4, possibly provide more space for players to accelerate and reach higher speeds.

Even when the available space is standardized (e.g., 100 m^2), larger formats like 4v4 played in rectangular spaces offer more open areas for longitudinal running, driven by the need to increase the field size to accommodate the relative area per player [11]. This can explain why, at speeds above 18 km/h, the 4v4 format accumulated the greatest distances, followed by the 3v3 format. Larger play-



▶ **Fig. 3** Comparisons between 2v2 and 4v4 formats based on player-to-area ratios. CI, confidence interval.



▶ **Fig. 4** Comparisons between 3v3 and 4v4 formats based on player-to-area ratios. CI, confidence interval.

ing areas provide greater opportunities to reach higher speeds, particularly during transition phases, such as counter-attacks or defensive recoveries [110]. In very congested spaces, such as those in 2v2, the available space may not allow for counter attacks or longitudinal exploration. However, this does not imply that the frequency of accelerations in 2v2 is necessarily higher than in other formats, as meta-analyses of acceleration frequencies across the different formats showed no significant differences.

The greater speeds and distances covered in 4v4 may be explained by the increased opportunity for players to explore more space, in line with previous studies on tactical collective behavior [111], which shows greater exploration of space and between-players distances in larger playing areas. This exploration is likely influenced by the affordances provided by the field layout and more structured positioning [112]. Thus, even at moderate speeds (7–13 and 13–17.8 km/h), the 4v4 format showed significantly greater distances covered compared to 3v3 and 2v2. It is expected that fundamental principles of play, particularly those related to mobility (e.g., finding space and creating passing lanes in depth near the last defender) and spatial awareness (e.g., stretching the field by occupying wide areas to disrupt the opponents' defensive focus), can also partially justify the results. These principles may help

▶ **Table 8** GRADE assessments for RPE, blood lactate and ball possession for the comparisons available between 2v2 and 4v4.

Outcome	*k (n)*	RoBAN-2	Indirectness	Risk of publication bias	Inconsistency	Imprecision	Certainty of evidence
Rating of perceived exertion							
2v2 vs. 3v3	15 (258)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by one level (I^2 = 74.3 %).	Downgrade by one level (< 800 participants). Favours 2v2.	⊕
2v2 vs. 4v4	18 (330)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 87.7 %).	Downgrade by one level (< 800 participants). Favours 2v2.	⊕
3v3 vs. 4v4	17 (281)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by one level (I^2 = 71.0 %).	Downgrade by one level (< 800 participants). Favours 3v3.	⊕
BLa concentration							
2v2 vs. 3v3	13 (224)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by 1 level (I^2 = 67.9 %).	Downgrade by one level (< 800 participants). Favours 2v2.	⊕
2v2 vs. 4v4	12 (209)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 90.3 %).	Downgrade by one level (< 800 participants). Favours 2v2.	⊕
3v3 vs. 4v4	17 (284)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 85.2 %).	Downgrade by one level (< 800 participants). Favours 3v3.	⊕
Total distance covered							
2v2 vs. 3v3	12 (198)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 95.6 %).	Downgrade by one level (< 800 participants). Favours 3v3.	⊕
2v2 vs. 4v4	13 (214)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 96.5 %).	Downgrade by one level (< 800 participants). Favours 4v4.	⊕
3v3 vs. 4v4	14 (282)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by two levels (I^2 = 93.1 %).	Downgrade by one level (< 800 participants). Favours 4v4.	⊕
Ball possession							
2v2 vs. 3v3	3 (60)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by one level (I^2 = 46.9 %).	Downgrade by two levels (< 800 participants and no clear direction of effects).	⊕
2v2 vs. 4v4	3 (60)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	Downgrade by one level (I^2 = 63.9 %).	Downgrade by one level (< 800 participants). Favours 2v2.	⊕
3v3 vs. 4v4	3 (60)	Downgrade by two levels (high risk).	No downgrading (low by default).	–	No downgrading (I^2 = 0.0 %).	Downgrade by one level (< 800 participants). Favours 3v3.	⊕

Abbreviations: BLa, blood lactate; CI, confidence interval; ES, effect size; RoB, risk of bias; RPE, rating of perceived exertion.

Note: Rules for the judgment: (i) All studies were non-randomized and therefore started at low certainty of evidence. Following the GRADE guidelines (see section 2.9 of the manuscript), they could be upgraded based on the presence of substantial ESs, effective control of potential confounders and evidence of a dose–response relationship, "if and only if", there were no reasons for downgrading (the five dimensions presented in this table). (ii) RoBAN 2: the risk of bias in studies: downgrading by one level in the presence of moderate RoB and by two levels in the presence of high RoB. (iii) Indirectness: low by default (see section 2.9). (iv) Risk of publication: no downgrading due to complex interpretation of existing tests [27]. (v) Inconsistency: downgraded by one level if I^2 = 25–75 % and by two levels if I^2 > 75 %. (vi) Imprecision: downgraded by one level if n < 800 (< 400 per group)[33] or effect direction unclear (95 % CIs crossing zero), or by two levels if both occurred.

⊕: very low certainty of evidence. ⊕⊕: low certainty of evidence.

explain why, in such approaches, the distance covered is greater, as the locomotor demands associated with exploring the field are increased. However, such multidimensional analysis often fails to take place, and future research should delve deeper into linking the emergence of tactical behaviors with the actual physical demands.

Interestingly, comparisons between 3v3 to 6v6 and 5v5 to 9v9 did not reveal significant differences, possibly because there is a threshold of space where the amount of available space is sufficient for running and covering distances, regardless of tactical factors. Furthermore, as the number of players increases within each for-

mat, player engagement may initially rise but eventually stabilize. This stabilization could be driven by more structured participation, with players assuming more defined positions on the field, which in turn affects the associated demands. After a certain level of defensive and attacking structures, influenced by the number of players, the collective behavior may shift. This can lead to more unit-based tactical behavior [113] which then requires exploiting boundaries or the depth of the opponent's defensive line. Such behavior may result in similar locomotor demands across these formats. However, studies integrating tactical analysis with physical demands are still limited, and a better understanding of how tactical modifications impact locomotor demands is needed. Future studies should aim to clearly describe the tactical changes that emerge from different formats and their subsequent effects on locomotor requirements.

Technical outcomes

By altering the number of players involved in the games, significant variations in technical actions are expected due to the impact on both individual and collective tactical behaviors. It was observed that both the 2v2 and 3v3 formats promoted ball possession more effectively than the 4v4 format [69–71], although no significant differences were found between the 2v2 and 3v3 formats. In smaller formats, such as 2v2 and 3v3, players may assume more diverse roles (both offensive and defensive) and are required to make decisions with less time due to increased pressure and reduced space [113]. This might promote better ball possession, as players are more engaged in the action while in 4v4, the dynamics may change, and players might be less involved or spread out, leading to more team-wide defensive coverage, which can affect individual ball possession skills. Interestingly, despite the increase in ball possession, no significant increase in ball losses was observed, as the number of lost balls remained similar and non-significant across the 2v2, 3v3, and 4v4 formats [58, 69, 103, 104].

On the other hand, the number of completed passes was significantly higher in the 3v3 format compared to the 2v2 format [69–71], which is somewhat surprising given that the smaller format might be expected to promote greater participation. In this context, the third attacker provides an extra passing option, which appears to increase the defensive challenge, as defenders must simultaneously monitor two attacking players rather than just one. These factors likely contribute to the easier ball circulation and therefore a higher number of completed passes in the 3v3 format. Interestingly, no significant differences were found between the other comparisons (2v2, 3v3, and 4v4). Despite these findings, completed passes are often analyzed in absolute terms rather than as a percentage, which can potentially confound the evidence. As the total number of passes increases, the number of completed passes may also rise, without necessarily maintaining a higher or lower completion percentage in comparison to absolute terms. Therefore, further research is needed, as factors such as time pressure, the increased number of defenders, or limited space can affect the overall efficacy of passes. Future studies should focus on standardizing the presentation of these values to provide a clearer understanding of pass efficacy, as well as other variables such as shots and dribbles.

Considering dribbling opportunities, only one comparison was possible for the meta-analysis, between 3v3 and 4v4 formats. The 3v3 format showed a significantly greater number of dribbles, with a large ES [37, 65, 74]. In the 3v3 format, the increased opportunities for engaging in duels and proximity to attacking zones, combined with a more varied participation in the game, may lead players to feel more frequently involved in attempts to break through the opponent's first defensive line [65]. Additionally, the possibly less structured defense in 3v3 may allow for more offensive freedom, without the defensive coverage and more structured balance seen in larger formats [113].

Finally, the meta-analyses focusing on successful shots revealed a significantly higher number of these actions in the 3v3 format compared to the 5v5 and 6v6 formats [41, 42, 59, 61, 65, 80, 91]. This difference could be attributed to several factors, including the closer proximity to the goals in 3v3, as the smaller playing space reduces the distance between players and the target. Additionally, the increased frequency of offensive situations in 3v3 might encourage more shooting opportunities. Another possible explanation is that the fewer number of players may lead to less defensive coverage, providing attacking players with more space and time to take shots. Future research could explore the role of defensive pressure, and how player positioning and tactics evolve in sided games, further influencing shooting success.

Interestingly, despite evidence showing that different-sided games impact the frequency of technical actions, no experimental studies have explored how various formats, applied continuously over a given period of time, affect skill development. This represents a crucial area for future research, particularly for coaches working with young and developing players, where the findings could have significant practical applications.

Limitations and future research

Conducting a systematic review with meta-analysis on the topic of play formats presents significant challenges due to the numerous concurrent factors that can influence the results. For instance, larger formats often involve larger playing spaces, even when the standardized space per player and the length-to-width ratio are controlled. When these conditions are maintained, the overall playing area increases, providing more opportunities for tactical variations. This, in turn, leads to differing physical demands. For example, in a larger space (e.g., 30 × 20; 150 m^2 per player), a sided format like 2v2 might not provide the necessary opportunities for sprinting (because of the longitudinal space available), while a format like 8v8 in a smaller space (e.g., 50 × 30; ~94 m^2 per player) may still offer sufficient longitudinal space for players to achieve sprinting. These complexities reduce the certainty of the evidence, making it difficult to draw a completely unbiased conclusion about the effects of different formats. This is an inherent challenge when treating the play format as an independent variable, as it often leads to different responses that reflect the dynamic, real-world challenges coaches face daily.

In our study, it was challenging to summarize the findings due to the inherent differences in comparisons across studies. For instance, in some comparisons between formats, the pitch-to-field area ratio was not maintained, while only 32 of the included studies

adhered to this ratio. Even within these studies, some maintained a 70 m^2 ratio, while others used 100 m^2 for the same formats. This leads to two major sources of variation: first, even when the player-to-area ratio is maintained, the size of the field can affect locomotor opportunities, as a larger field inherently changes movement dynamics; second, even when comparing the same formats in two different studies, variations in the player-to-area ratio introduce inherent heterogeneity. These factors should be clearly acknowledged when interpreting the results, and coaches must be aware of these variations to avoid drawing overly confident conclusions from the evidence.

Furthermore, research on sided games faces a higher RoB due to confounding factors, particularly when randomization is not employed or when the sequence of play is randomized. External influences, such as the time of day or specific circumstances, can introduce heterogeneity into the results, complicating the interpretation of findings. The selection of outcomes also presents a potential RoB. Many studies in this field were not pre-registered, allowing for the addition or removal of outcomes after the experiment. This flexibility increases the risk of data manipulation or selective reporting.

The presence of a high RoB due to confounding, incomplete outcome data, and selective outcome reporting can also affect the interpretation of the ultimate findings. Confounding can distort the true relationship between variables, leading to misleading conclusions about the effectiveness or impact of an intervention. Incomplete outcome data further complicates the situation, as missing data may create an inaccurate representation of the results, potentially skewing conclusions. Selective outcome reporting, where only certain results are reported or emphasized, increases the risk of reporting bias, thus undermining the reliability of the evidence. To mitigate these issues in future research, pre-registration of study protocols would provide a clear plan for data collection and outcome reporting, reducing the likelihood of selective reporting. Additionally, ensuring clearer and more transparent reporting of randomisation processes would help strengthen the validity of findings, as it would reduce the RoB introduced through inadequate or unclear randomisation methods.

Another challenge arises when assessing outcomes in studies with small sample sizes, such as when only 8 or 10 players are chosen from a team of around 20. The absence of clear criteria for player selection can create bias, as it is unclear why certain players were chosen over others. This lack of transparency should be addressed in this research.

Additionally, while some studies provide detailed context about the experiments (e.g., season timing, day of the week, recovery period, environmental conditions, and time of day), others lack similar information. Given that acute responses to sided games can be influenced by these external, often uncontrolled factors, it is important to avoid making definitive conclusions based on the current study. Furthermore, future studies should consider these contextual variables as potential sources of variation, incorporating them into hierarchical or multivariate models to better control for their impact in comparisons between formats.

To improve the rigor and replicability of future research, it is essential to strengthen the methodological approach and reporting standards. This includes clearly defining the criteria for participant selection, registering studies a priori, and providing a detailed context about the experimental setup. Additionally, confounding factors can be better controlled through multicenter studies with larger, more diverse populations and by conducting longitudinal studies that assess how different circumstances might influence the results over time. These measures will help enhance the quality of evidence and provide more meaningful insights into the effects of play formats.

Practical applications

Any definitive conclusions should be avoided based on these results, as the certainty of evidence is low to very low. While the meta-analysis indicates statistical trends, definitive evidence is not available, and practical applications should be approached with caution, especially considering the previously discussed limitations. Nevertheless, if practical recommendations can be drawn from the current meta-analysis, it suggests that formats smaller than 4v4 are the most physiologically demanding. These formats may be relevant for targeting aerobic power zones. Blood lactate levels and perceived efforts are clearly higher in smaller formats such as 2v2 and 3v3, while HR measures do not significantly differ from 1v1 to 4v4 formats. However, after 5v5, HR intensity trends tend to decline and are significantly lower compared to 4v4. Therefore, if the goal is to introduce an appropriate cardiorespiratory stimulus with more intense demands, smaller formats such as 4v4 or even smaller may be the most suitable choices for such training target. However, the heterogeneity among the included study designs, as well as the influence of other concurrent task constraints (e.g., player-to-area ratio and type of scoring) or external factors (e.g., environment and timing within the season), should be carefully considered. These factors may moderate any definitive interpretations derived from the statistical trends identified.

On the other hand, coaches may opt for slightly larger formats if the goal is to emphasize locomotor demands. In smaller formats, such as 2v2 or 3v3, the distances covered tend to be significantly shorter compared to 4v4 or larger formats. Additionally, this decision is influenced by the available space, which interacts with the chosen format to shape locomotor demands. For example, the meta-analyses indicate that 4v4 games generally require significantly greater locomotor effort compared to 2v2 or 3v3 across different speed thresholds. Conversely, when comparing LSGs, such as 9v9 or 10v10, the distance covered is often greater in mid-sized formats like 5v5 or 6v6. Therefore, to ensure an appropriate accumulation of the distance covered across different thresholds, formats ranging between 4v4 and 6v6 may be particularly interesting. However, it is also important to be cautious with these interpretations, as factors such as field dimensions, field configuration (including the player-to-area ratio), and area per player can significantly influence the results. Based on this, coaches should carefully evaluate the use of different formats in relation to field dimensions and avoid drawing definitive conclusions due to the low certainty of evidence caused by the heterogeneity of the studies included.

Considering the training focus on technical elements, the meta-analysis shows that, when comparing 3v3 with 5v5 or 6v6, the 3v3 format increases the number of successful shots made by players. Additionally, 3v3 significantly promotes more dribbles than 4v4. Ball possession is also significantly greater in 2v2 and 3v3 for-

mats compared to 4v4. In terms of completed passes, 3v3 outperforms 4v4. In summary, the 3v3 format appears to be particularly effective for technical training, especially in terms of ball possession, without increasing the number of unsuccessful passes. It also encourages a higher number of shot-related actions, making it a highly relevant format for introducing these elements as the primary focus. However, coaches should treat this information as statistical trends rather than definitive evidence, as technical actions were primarily collected in absolute numbers rather than in terms of accuracy. This could lead to differences in efficacy, potentially conflicting with the expected results. Additionally, the available area may impose constraints, as time pressure can vary and impact the results. Therefore, caution is needed when translating these statistical terms into practical applications.

Conclusions

Our systematic review with meta-analysis suggests that SSG formats (2v2 and 3v3) are the most physiologically demanding, with significantly higher perceived exertion and blood lactate levels compared to larger formats. While HR measures showed no clear trends, 3v3 exhibited greater intensity than 4v4 and 5v5. In terms of physical demands, 4v4 and 3v3 covered more distance than 2v2, although beyond 4v4, differences became less pronounced. Smaller formats also enhanced ball possession, with 3v3 promoting more successful shots and dribbles than larger formats. However, physical demands increased in mid-sized formats (4v4–6v6), making them suitable for targeting locomotor demands. All these findings should be interpreted with caution due to the low certainty of evidence, as well as the existing heterogeneity and RoB identified in the studies.

Funding

José Afonso, CIFI2D, is financed by the Portuguese Foundation for Science and Technology. 10.54499/UIDB/05913/2020.

Conflict of Interest

The authors declare that they have no conflict of interest.

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: 199–220

[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: 154–161

[3] Clemente FM, Afonso J, Silva RM et al. Contemporary practices of Portuguese and Brazilian soccer coaches in designing and applying small-sided games. Biol Sport 2024; 41: 185–199. DOI: 10.5114/biolsport.2024.132985

[4] Ometto L, Vasconcellos FV, Cunha FA et al. How manipulating task constraints in small-sided and conditioned games shapes emergence of individual and collective tactical behaviours in football: A systematic review. Int J Sports Sci Coach 2018; 13: 1200–1214

[5] Clemente FM, Moran J, Ramirez-Campillo R, Beato M, Afonso J. Endurance performance adaptations between SSG and HIIT in soccer players: A meta-analysis. Int J Sports Med 2024; 45: 183–210. DOI: 10.1055/a-2171-3255

[6] Bujalance-Moreno P, Latorre-Román PÁ, García-Pinillos F. A systematic review on small-sided games in football players: Acute and chronic adaptations. J Sports Sci 2019; 37: 921–949

[7] Clemente FM, Afonso J, Sarmento H. Small-sided games: An umbrella review of systematic reviews and meta-analyses. PLoS One 2021; 16: e0247067

[8] Francesco Sgrò S, Salvatore Bracco B, Salvatore Pignato P Mario Lipoma. Small-Sided Games and Technical Skills in Soccer Training: Systematic Review and Implications for Sport and Physical Education Practitioners. J Sports Sci 2018; 6: 9–19. DOI: 10.17265/2332-7839/2018.01.002

[9] Ferreira-Ruiz Á, García-Banderas F, Martín-Tamayo I. Systematic Review: Technical-Tactical Behaviour in Small-Sided Games in Men's Football. Apunts Educ Fís Deportes 2022; 148: 42–61

[10] DE Dios-Álvarez V, Lorenzo-Martínez M, Padrón-Cabo A, Rey E. Small-sided games in female soccer players: a systematic review. J Sports Med Phys Fitness 2022; 62: 1474–1480

[11] Castagna C, Francini L, Póvoas SCA, D'Ottavio S. Long-sprint abilities in soccer: Ball versus running drills. Int J Sports Physiol Perform 2017; 12: 1256–1263

[12] Hill-Haas S, Coutts A, Rowsell G, Dawson B. Variability of acute physiological responses and performance profiles of youth soccer players in small-sided games. J Sci Med Sport 2008; 11: 487–490

[13] Owen AL, Wong DP, Paul D, Dellal A. Physical and technical comparisons between various-sided games within professional soccer. Int J Sports Med 2014; 35: 286–292

[14] Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71

[15] Rechenchosky L, Menegassi VM, Jaime MO et al. Scoping review of tests to assess tactical knowledge and tactical performance of young soccer players. J Sports Sci 2021; 39: 2051–2067

[16] McKay AKA, Stellingwerff T, Smith ES et al. Defining Training and Performance Caliber: A Participant Classification Framework. Int J Sports Physiol Perform 2022; 17: 317–331

[17] Wojtyniak JG, Britz H, Selzer D, Schwab M, Lehr T. Data Digitizing: Accurate and Precise Data Extraction for Quantitative Systems Pharmacology and Physiologically-Based Pharmacokinetic Modeling. CPT Pharmacometrics Syst Pharmacol 2020; 9: 322–331

[18] Seo H-J, Kim SY, Lee YJ, Park JE. RoBANS 2: A Revised Risk of Bias Assessment Tool for Nonrandomized Studies of Interventions. Korean J Fam Med 2023; 44: 249–260

[19] Valentine JC, Pigott TD, Rothstein HR. How Many Studies Do You Need? J Educ Behav Stat 2010; 35: 215–247

[20] Abt G, Boreham C, Davison G et al. Power, precision, and sample size estimation in sport and exercise science research. J Sports Sci 2020; 38: 1933–1935

[21] Zouhal H, Hammami A, Tijani JM et al. Effects of Small-Sided Soccer Games on Physical Fitness, Physiological Responses, and Health Indices in Untrained Individuals and Clinical Populations: A Systematic Review. Sports Med 2020; 50: 987–1007

[22] López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses. Res Synth Methods 2018; 9: 336–351

[23] Deeks JJ, Higgins JP, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions.The Cochrane Collaboration; 2008: pp 243–296

[24] Kontopantelis E, Springate DA, Reeves D. A Re-Analysis of the Cochrane Library Data: The Dangers of Unobserved Heterogeneity in Meta-Analyses. PLoS One 2013; 8: e69930

[25] Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive Statistics for Studies in Sports Medicine and Exercise Science. Med Sci Sports Exerc 2009; 41: 3–13

[26] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315: 629–634

[27] Afonso J, Ramirez-Campillo R, Clemente FM, Büttner FC., Andrade R. The Perils of Misinterpreting and Misusing "Publication Bias" in Meta-analyses: An Education Review on Funnel Plot-Based Methods. Sports Med 2024; 54: 257–269

[28] Duval S, Tweedie R. Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. Biometrics 2000; 56: 455–463

[29] Zhang Y, Alonso-Coello P, Guyatt GH et al. GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences—Risk of bias and indirectness. J Clin Epidemiol 2019; 111: 94–104

[30] Zhang Y, Coello PA, Guyatt GH et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences—inconsistency, imprecision, and other domains. J Clin Epidemiol 2019; 111: 83–93

[31] Guyatt GH, Oxman AD, Sultan S et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 2011; 64: 1311–1316

[32] Balshem H, Helfand M, Schünemann HJ et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011; 64: 401–406

[33] Guyatt G, Oxman AD, Akl EA et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011; 64: 383–394

[34] Schünemann HJ, Cuello C, Akl EA et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol 2019; 111: 105–114

[35] Guyatt G, Oxman AD, Kunz R et al. Corrigendum to GRADE guidelines 6. Rating the quality of evidence-imprecision. J Clin Epidemiol 2011;64:1283–1293. J Clin Epidemiol 2021; 137: 265

[36] Aasgaard M, Kilding AE. Does Man Marking Influence Running Outputs and Intensity During Small-Sided Soccer Games? J Strength Cond Res 2020; 34: 3266–3274

[37] Abrantes CI, Nunes MI, Maças VM, Leite NM, Sampaio JE. Effects of the number of players and game type constraints on heart rate, rating of perceived exertion, and technical actions of small-sided soccer games. J Strength Cond Res 2012; 26: 976–981

[38] Ade JD, Harley JA, Bradley PS. Physiological response, time-motion characteristics, and reproducibility of various speed-endurance drills in elite youth soccer players: small-sided games versus generic running. Int J Sports Physiol Perform 2014; 9: 471–479

[39] Aguiar MV, Botelho GM, Gonçalves BS, Sampaio JE. Physiological responses and activity profiles of football small-sided games. J Strength Cond Res 2013; 27: 1287–1294

[40] Alashti M, Shirvani H, Sabzevari Rad R, Noori Ordeghan A. Effects of different small-sided games on blood lactate and GH/IGF-1 axis responses in young soccer players. German J Exerc Sport Res 2021; 51: 86–93

[41] Alcântara CH, Teixeira AS, Teixeira RM et al. Manipulation of number of players and bouts duration in small-sided games in youth soccer players. Sport Sci Health 2021; 17: 597–605

[42] Almeida CH, Ferreira AP, Volossovitch A. Offensive sequences in youth soccer: effects of experience and small-sided games. J Hum Kinet 2013; 36: 97–106

[43] Arslan E, Alemdaroglu U, Koklu Y, Hazir T, Muniroglu S, Karakoc B. Effects of Passive and Active Rest on Physiological Responses and Time Motion Characteristics in Different Small Sided Soccer Games. J Hum Kinet 2017; 60: 123–132

[44] Aşçı A. Heart Rate Responses during Small Sided Games and Official Match-Play in Soccer. Sports 2016; 4: 31. DOI: 10.3390/sports4020031

[45] Asian-Clemente JA, Rabano-Muñoz A, Núñez FJ, Suarez-Arrones L. External and internal load during small-sided games in soccer: use or not floaters. J Sports Med Phys Fitness 2022; 62: 301–307

[46] Asín-Izquierdo I, Gutiérrez-García L, Galiano C. Application of technology for the analysis of Small-Sided Games in football. From complexity to chaos in training design: Reference to number of players, playing space, orientation, time distribution, directionality with goalkeepers, and feedback. Proc Inst Mech Eng P J Sport Eng Technol 2023; 238: 117–125. DOI: 10.1177/17543371231175946

[47] Beato M, Bertinato L, Schena F. High volume training with small-sided games affects technical demands in football: a descriptive study. Sport Sci Health 2014; 10: 219–223

[48] Beato M, Vicens-Bordas J, Peña J, Costin AJ. Training load comparison between small, medium, and large-sided games in professional football. Front Sports Act Living 2023; 5: 1165242

[49] Beenham M, Barron DJ, Fry J, Hurst HH, Figueirdo A, Atkins S. A Comparison of GPS Workload Demands in Match Play and Small-Sided Games by the Positional Role in Youth Soccer. J Hum Kinet 2017; 57: 129–137

[50] Bergmann F, Braksiek M, Meier C. The influence of different game formats on technical actions and playing time parameters - A study with under-7 and under-9 soccer players in a competitive context. Int J Sports Sci Coach 2022; 17: 1089–1100

[51] Bergmann F, Meier C, Braksiek M. Involvement and performance of U9 soccer players in 7v7 and 5v5 matches during competition. German J Exerc Sport Res 2022; 52: 125–134

[52] Brandes M, Heitmann A, Müller L. Physical responses of different small-sided game formats in elite youth soccer players. J Strength Cond Res 2012; 26: 1353–1360

[53] Bravo-Sánchez A, Abián-Vicén J, Abián P. Analysis of the physical and technical differences between 7-a-side and 8-a-side game modalities in official under 12 soccer matches. Int J Perform Anal Sport 2017; 17: 545–554

[54] Brito Â, Roriz P, Duarte R, Garganta J. Match-running performance of young soccer players in different game formats. Int J Perform Anal Sport 2018; 18: 410–422

[55] Casamichana D, Román-Quintana JS, Castellano J, Calleja-González J. Influence of the Type of Marking and the Number of Players on Physiological and Physical Demands During Sided Games in Soccer. J Hum Kinet 2015; 47: 259–268

[56] Castellano J, Casamichana D, Dellal A. Influence of game format and number of players on heart rate responses and physical demands in small-sided soccer games. J Strength Cond Res 2013; 27: 1295–1303

[57] Castellano J, Puente A, Echeazarra I, Usabiaga O, Casamichana D. Number of Players and Relative Pitch Area per Player: Comparing Their Influence on Heart Rate and Physical Demands in Under-12 and Under-13 Football Players. PLoS One 2016; 11: e0127505

[58] Clemente FM. Greater Psychophysiological Intensities in Conditioned Games May Impair Technical Performance: An Exploratory Study in Youth Male Soccer Players. Int J Sports Physiol Perform 2024; 19: 576–584

[59] Clemente FM, Chen YS, Bezerra JP, Guiomar J, Lima R. Between-format differences and variability of technical actions during small-sided soccer games played by young players. Human Movement 2018; 2018: 114–120

[60] Clemente FM, Praça GM, Bredt S, van der Linden C, Serra-Olivares J. External Load Variations Between Medium- and Large-Sided Soccer Games: Ball Possession Games vs Regular Games with Small Goals. J Hum Kinet 2019; 70: 191–198

[61] Clemente FM, Sarmento H, Costa IT, Enes AR, Lima R. Variability of Technical Actions During Small-Sided Games in Young Soccer Players. J Hum Kinet 2019; 69: 201–212

[62] Clemente FM, Sarmento H, Rabbani A, Van Der Linden C, Kargarfard M, Costa IT. Variations of external load variables between medium- and large-sided soccer games in professional players. Res Sports Med 2019; 27: 50–59

[63] Clemente FM, Silva AF, Kawczyński A et al. Physiological and locomotor demands during small-sided games are related to match demands and physical fitness? A study conducted on youth soccer players. BMC Sports Sci Med Rehabil 2022; 14: 138

[64] Costa TS, Rossato M, Rodrigues O et al. The impact of different game formats on players' and team performance in youth soccer competitions. Int J Sports Sci Coach 2024; 19: 2016–2023

[65] da Silva CD, Impellizzeri FM, Natali AJ et al. Exercise intensity and technical demands of small-sided games in young Brazilian soccer players: effect of number of players, maturation, and reliability. J Strength Cond Res 2011; 25: 2746–2751

[66] Dalen T, Sandmæl S, Stevens TGA et al. Differences in Acceleration and High-Intensity Activities Between Small-Sided Games and Peak Periods of Official Matches in Elite Soccer Players. J Strength Cond Res 2021; 35: 2018–2024

[67] Dellal A, Chamari K, Owen AL et al. Influence of technical instructions on the physiological and physical demands of small-sided soccer games. Eur J Sport Sci 2011; 11: 341–346

[68] Dellal A, Chamari K, Pintus A, Girard O, Cotte T, Keller D. Heart rate responses during small-sided games and short intermittent running training in elite soccer players: a comparative study. J Strength Cond Res 2008; 22: 1449–1457

[69] Dellal A, Drust B, Lago-Penas C. Variation of Activity Demands in Small-Sided Soccer Games. Int J Sports Med 2012; 33: 370–375

[70] Dellal A, Hill-Haas S, Lago-Penas C, Chamari K. Small-sided games in soccer: amateur vs. professional players' physiological responses, physical, and technical activities. J Strength Cond Res 2011; 25: 2371–2381

[71] Dellal A, Jannault R, Lopez-Segovia M, Pialoux V. Influence of the Numbers of Players in the Heart Rate Responses of Youth Soccer Players Within 2 vs. 2, 3 vs. 3 and 4 vs. 4 Small-sided Games. J Hum Kinet 2011; 28: 107–114

[72] Dimitriadis Y, Michailidis Y, Mandroukas A et al. Internal and external load of youth soccer players during small-sided games. Trends Sport Sci 2022; 29: 171–181

[73] Edis Ç, Vural F, Vurgun H. The Importance of Postural Control in Relation to Technical Abilities in Small-Sided Soccer Games. J Hum Kinet 2016; 53: 51–61

[74] Bekris E, Eleftherios M, Aris S et al. Supernumerary in small sided games 3vs3 & 4vs4. J Phys Educ Sport 2012; 12: 398–406

[75] Farhani Z, Hammami R, Gene-Morales J et al. Bout duration and number of players of soccer small-sided games affect perceived enjoyment, physiological responses, and technical-tactical performance. Sci Med Footb 2022; 6: 503–510

[76] Febré R, Chirosa LJ, Casamichana D et al. Influencia de la densidad de jugadores sobre la frecuencia cardíaca y respuestas técnicas en jóvenes jugadores de fútbol. [Influence of the density of players on their heart rate and its technical implications on young football players]. Rev Int Cienc Deporte 2015; 11: 116–128

[77] García-Angulo A, Palao JM, Giménez-Egido JM, García-Angulo FJ, Ortega-Toro E. Effect of the Modification of the Number of Players, the Size of the Goal, and the Size of the Field in Competition on the Play Actions in U-12 Male Football. Int J Environ Res Public Health 2020; 17: 518. DOI: 10.3390/ijerph17020518

[78] Giménez JV, Del-Coso J, Leicht AS, Gomez MÁ. Comparison of the movement patterns between small- and large-sided game training and competition in professional soccer players. J Sports Med Phys Fitness 2018; 58: 1383–1389

[79] Halouani J, Chtourou H, Dellal A, Chaouachi A, Chamari K. Soccer small-sided games in young players: rule modification to induce higher physiological responses. Biol Sport 2017; 34: 163–168

[80] Harrison CB, Kilding AE, Gill ND, Kinugasa T. Small-sided games for young athletes: is game specificity influential? J Sports Sci 2014; 32: 336–344

[81] Hill-Haas SV, Dawson BT, Coutts AJ, Rowsell GJ. Physiological responses and time–motion characteristics of various small-sided soccer games in youth players. J Sports Sci 2009; 27: 1–8

[82] Jastrzębski Z, Radzimiński Ł. Individual vs general time-motion analysis and physiological response in 4 vs 4 and 5 vs 5 small-sided soccer games. Int J Perform Anal Sport 2015; 15: 397–410

[83] Köklü Y, Alemdaroğlu U, Cihan H, Wong DP. Effects of bout duration on players' internal and external loads during small-sided games in young soccer players. Int J Sports Physiol Perform 2017; 12: 1370–1377

[84] Köklü Y. A comparison of physiological responses to various intermittent and continuous small-sided games in young soccer players. J Hum Kinet 2012; 31: 89–96

[85] Köklü Y, Albayrak M, Keysan H et al. Improvement of the physical conditioning of young soccer players by playing small-sided games on different pitch size - special reference to physiological responses. Kinesiology 2013; 45: 41–47

[86] Köklü Y, Alemdaroğlu U. Comparıson of the Heart Rate and Blood Lactate Responses of Different Small Sided Games in Young Soccer Players. Sports 2016; 4: 48

[87] Köklü Y, Aşçi A, Koçak FU, Alemdaroğlu U, Dündar U. Comparison of the physiological responses to different small-sided games in elite young soccer players. J Strength Cond Res 2011; 25: 1522–1528

[88] Köklü Y, Sert Ö, Alemdaroğlu U, Arslan Y. Comparison of the physiological responses and time-motion characteristics of young soccer players in small-sided games: The effect of goalkeeper. J Strength Cond Res 2015; 29: 964–971

[89] Köklü Y, Türkdoğan H, Bölükbaşı T, Alemdaroğlu U. Comparison of internal and external loads during different small-sided games in young female soccer players. Sci Sports 2024; 39: 260–266

[90] Little T, Williams AG. Measures of Exercise Intensity During Soccer Training Drills With Professional Soccer Players. J Strength Cond Res 2007; 21: 367–371

[91] Lex H, Simon M, Schwab S. Insights into the application of soccer-specific actions in established and new game forms of youth soccer. German J Exerc Sport Res 2022; 52: 168–172

[92] López-Fernández J, Sánchez-Sánchez J, García-Unanue J, Hernando E, Gallardo L. Physical and Physiological Responses of U-14, U-16, and U-18 Soccer Players on Different Small-Sided Games. Sports (Basel) 2020; 8: 66. DOI: 10.3390/sports8050066

[93] Makar P, Silva AF, Kawczyński A et al. Variability of peak speed and sprinting actions during the same small-sided games: within- and between-player variations inspected over four consecutive weeks. Biol Sport 2023; 40: 959–965

[94] Mara JK, Thompson KG, Pumpa KL. Physical and Physiological Characteristics of Various-Sided Games in Elite Women's SoccerPhysical and Physiological Characteristics of Various-Sided Games in Elite Women's Soccer. Int J Sports Physiol Perform 2016; 11: 953–958

[95] Martin-Garcia A, Castellano J, Diaz AG, Cos F, Casamichana D. Positional demands for various-sided games with goalkeepers according to the most demanding passages of match play in football. Biol Sport 2019; 36: 171–180

[96] Owen AL, Wong del P, 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: 2104–2110

[97] Papadopoulos EK, Tsentidou G, Metaxas TI et al. The effect of pitch dimensions and players' format on heart load and external load in semi-professional soccer players. Trends Sport Sci 2023; 30: 175–186

[98] Pinheiro GS, Chiari Quintão R, Nascimento VB, Claudino JG, Alves AL, Teoldo da Costa I, Teoldo da Costa V. Small-sided games do not replicate all external and internal loads of a football match-play during pre-season: A case study. Int J Sports Sci Coach 2023; 18: 152–159

[99] Rampinini E, Impellizzeri FM, Castagna C, Abt G, Chamari K, Sassi A, Marcora SM. Factors influencing physiological responses to small-sided soccer games. J Sports Sci 2007; 25: 659–666

[100] Rebelo AN, Silva P, Rago V, Barreira D, Krustrup P. Differences in strength and speed demands between 4v4 and 8v8 small-sided football games. J Sports Sci 2016; 34: 2246–2254

[101] Savolainen EHJ, Vänttinen T, Peltonen J, Ihalainen JK, Walker S. Average demands and most demanding passages of national-level female soccer matches: do small- and large-sided games replicate match demands? Front Sports Act Living 2023; 5: 1236112. DOI: 10.3389/fspor.2023.1236112

[102] Silva AF, González-Fernández FT, Aquino R, Akyildiz Z, Vieira LP, Yıldız M, Birlik S, Nobari H, Praça G, Clemente FM. Analyzing the within and between Players Variability of Heart Rate and Locomotor Responses in Small-Sided Soccer Games Performed Repeatedly over a Week. Healthcare (Basel) 2022; 10: 1412. DOI: 10.3390/healthcare10081412

[103] Wang T, Xue T, He J. Do Different Playing Formats and Aerobic Capacity Influence Variances in Psychophysiological Demands and Technical Performance in Small-Sided Games? A Study among Under-17 Male Soccer Players. J Sports Sci Med 2024; 23: 619–627

[104] Yılmaz O, Soylu Y. A Comparative Study of the Effects of Small-Sided Game Formats on Internal Load and Technical Responses in Soccer. Pamukkale J Sport Sci 2024; 15: 416–431

[105] Younesi S, Rabbani A, Manuel Clemente F, Sarmento H, Figueiredo A. Session-to-session variations of internal load during different small-sided games: a study in professional soccer players. Res Sports Med 2021; 29: 462–474

[106] Younesi S, Rabbani A, Clemente FM, Sarmento H, Figueiredo AJ. Session-to-session variations in external load measures during small-sided games in professional soccer players. Biol Sport 2021; 38: 185–193

[107] Köklü Y, Alemdaroğlu U, Cihan H, Wong DP. Effects of Bout Duration on Players' Internal and External Loads During Small-Sided Games in Young Soccer Players. Int J Sports Physiol Perform 2017; 12: 1370–1377

[108] Coutts AJ, Rampinini E, Marcora SM, Castagna C, Impellizzeri FM. Heart rate and blood lactate correlates of perceived exertion during small-sided soccer games. J Sci Med Sport 2009; 12: 79–84

[109] Clemente FM. The threats of small-sided soccer games: A discussion about their differences with the match external load demands and their variability levels. Strength Cond J 2020; 42: 100–105

[110] Fradua L, Zubillaga A, Caro O, Iván Fernández-García A, Ruiz-Ruiz C, Tenga A. Designing small-sided games for training tactical aspects in soccer: Extrapolating pitch sizes from full-size professional matches. J Sports Sci 2013; 31: 573–581

[111] Aguiar M, Gonçalves B, Botelho G, Lemmink K, Sampaio J. Footballers' movement behaviour during 2-, 3-, 4- and 5-a-side small-sided games. J Sports Sci 2015; 33: 1259–1266

[112] Olthof SBH, Frencken WGP, Lemmink KAPM. A match-derived relative pitch area facilitates the tactical representativeness of small-sided games for the official soccer match. J Strength Cond Res 2019; 33: 523–530

[113] Silva B, Garganta J, Santos R, Teoldo I. Comparing Tactical Behaviour of Soccer Players in 3 vs. 3 and 6 vs. 6 Small-Sided Games. J Hum Kinet 2014; 41: 191–202