

Action-Quality rating systems in indoor volleyball: A scoping review of conceptual frameworks, analytical approaches, and positional applications

International Journal of Sports Science
& Coaching
1–13

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/17479541261415737

journals.sagepub.com/home/spo

Zeljko L Labit¹ and Kirsten N Spencer¹

Abstract

Objective: This scoping review aimed to systematically map and categorise action-quality rating systems used to evaluate individual skill execution in indoor volleyball, with practical relevance for performance analysts supporting elite-level athletes such as the New Zealand VolleyFerns.

Methods: Following the Joanna Briggs Institute and PRISMA-ScR guidelines, a comprehensive literature search was conducted across PubMed, SPORTDiscus, and Scopus. Eligible studies included those that developed, applied, or evaluated rating systems for technical skill execution in indoor volleyball from 2000 to 2024. Forty-four studies were included after structured screening and data extraction using CADIMA. Rating systems were classified by typology, analytical approach, application context, and positional specificity.

Results: Three typological categories emerged: ordinal-descriptive systems (64%), interval-probabilistic models (9%), and hybrid-emergent frameworks (27%). Analytical approaches ranged from basic descriptive statistics to probabilistic modelling, social network analysis, and machine learning. Most studies focused on elite-level female athletes and emphasised terminal actions (e.g., serve, attack), with growing attention to contextual and positional variation. Setter and libero actions were increasingly analysed through system-sensitive models reflecting match phase and tactical role.

Conclusion: Volleyball action-quality rating systems have evolved from static ordinal scales to complex, context-sensitive frameworks that better capture the dynamic relationship between skill execution and match context. However, challenges such as methodological heterogeneity, inconsistent terminology, and limited validation persist. This review provides foundational insights for developing robust, role-specific, and analytically sound rating systems to inform coaching and talent identification in elite volleyball.

Keywords

Gender, performance analysis, player evaluation, skill, system modelling, talent identification

Introduction

Performance analysis in volleyball has traditionally emphasised outcome-based indicators such as point differentials, win-loss ratios, or aggregate team statistics. However, over the past two decades, analysts have shifted their focus toward process-oriented evaluation. Specifically, the systematic assessment of how well players execute individual skills within the flow of match play. Central to this shift is the development of action-quality rating systems, which provide a structured method for coding, scoring, and interpreting player actions in context. These systems underpin modern match analysis by aiming to quantify not only what happened (e.g., a kill, an error) but also how effectively players performed each action relative to technical, tactical, or situational benchmarks. While the Fédération

Internationale de Volleyball (FIVB) popularised ordinal coding schemes as early as the 1970s, many of these initial models¹ were based on heuristic judgements of outcome success rather than validated performance metrics.² Since

Reviewers: Sotirios Driskos (National and Kapodistrian University of Athens, Greece)

¹Sports Performance Research Institute New Zealand, Auckland University of Technology, Auckland, New Zealand

Corresponding author:

Kirsten N Spencer, Sports Performance Research Institute New Zealand, Auckland University of Technology, 90 Akoranga Drive, Auckland, 1142, New Zealand.
Email: kspencer@aut.ac.nz

then, researchers and analysts have proposed a diverse array of rating frameworks, some maintaining descriptive ordinal categories (e.g., ‘perfect pass,’ ‘limited options’), others adopting interval or probabilistic models that estimate point-scoring probabilities based on Markov chains or Bayesian inference.^{3,4} These systems vary not only in scale structure (ordinal, interval, continuous) but also in their theoretical underpinnings, methodological rigour, and practical usability in coaching and scouting contexts.

Despite their widespread use, action-quality rating systems in volleyball are highly fragmented. Researchers often report rating frameworks with inconsistent terminology, a lack of justification, and little explanation of validation methods. This inconsistency makes it hard to compare systems and synthesise findings. For coaches, inconsistent rating logic makes performance reports tough to interpret, benchmarking difficult, and practical training decisions less clear. Some studies assess isolated skills like setting or serving, while others analyse whole game phases using combined skills. Some rating systems are backed by reliable observation tools, but others lack methodological transparency² and reduce coach confidence. This diversity presents both a challenge and an opportunity. It limits the creation of coach-relevant benchmarks for monitoring, selection, and tactical evaluation, but it also shows progress in performance analysis, recognising that action quality is tied to roles and match dynamics. Previous research accentuates the need to map and classify evaluation frameworks and clarify their use for skill assessment in real matches.⁵

Accordingly, the rationale for this scoping review gains further support from its applied relevance. Volleyball New Zealand (VNZ), through its elite women’s program (the VolleyFerns), is in the process of establishing a national performance baseline for its top 48 female players. A key component of this initiative involves using action-quality ratings to assess positional execution, team system effectiveness, and individual decision-making under pressure during the upcoming VolleyFerns Invitational (VFI) Tournament. However, selecting or developing appropriate rating instruments requires a comprehensive understanding of what systems exist, how researchers have implemented them in the literature, and which ones have demonstrated validity, reliability, or contextual adaptability. Without such a synthesis, performance analysts risk using tools that may lack conceptual clarity, empirical robustness, or practical alignment with coaching needs. Scoping reviews, as articulated by Peters et al.,⁶ are particularly well-suited to address such gaps. Unlike systematic reviews, which typically aim to synthesise findings around narrowly defined intervention effects, scoping reviews aim to map broad bodies of literature, identify key concepts, clarify definitions, and chart the range of evidence available across disciplines, populations, and contexts. This scoping review aims to systematically map and categorise action-quality rating systems used to assess individual skill execution in indoor volleyball, providing a critical

foundation for evidence-based performance evaluation frameworks that can enhance coaching effectiveness, inform talent identification processes, and optimise training methodologies in elite volleyball programs.

Methods

Protocol and registration

This scoping review was conducted following the methodological guidance outlined by the Joanna Briggs Institute (JBI) for scoping reviews⁶ and is reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines.⁷ The purpose of this methodological approach is to ensure transparency, replicability, and methodological rigour in the mapping of existing literature related to action-quality rating systems in volleyball. The protocol defined the population, concept, and context (PCC) parameters, inclusion and exclusion criteria, and analytical procedures for data extraction and synthesis. The protocol was not registered in an open-access repository.

Eligibility criteria

Studies were included in this review if they met the following criteria: (1) volleyball athletes (male or female), teams, or match contexts at any competitive level; (2) studies that described, applied, developed, or evaluated an action-quality rating system for assessing individual technical skill execution; and (3) indoor volleyball match or training environments; observational or analytical studies employing qualitative, quantitative, or mixed-methods frameworks. Additional inclusion criteria included: (1) full-text availability in English; (2) peer-reviewed journal articles, conference proceedings, or validated research instruments published between 2000 and 2024; (3) explicit description of rating system typology; and (4) discussion or application of performance evaluation, not merely frequency counts or descriptive statistics. Studies were excluded if they: (1) focused solely on male athletes, beach volleyball, physical testing, biomechanics, or health-related outcomes (e.g., injury, physiology); (2) did not apply or develop a skill-rating system; and (3) were purely narrative commentaries, instructional manuals, or abstracts without methods or results.

Information sources and search strategy

The literature search was conducted across the following electronic databases: PubMed/MEDLINE, SPORTDiscus, and Scopus. The search strategy was developed and conducted in accordance with the Peer Review of Electronic Search Strategies (PRESS) checklist to ensure its robustness and comprehensiveness.⁸ The search strategy encompasses both controlled vocabulary (e.g., MeSH terms) and free-text terms to capture all relevant literature. The key search terms were

'volleyball,' 'skill performance,' and 'rating systems,' along with their synonyms, related terms, and any appropriate variations. Database-specific syntaxes were adapted to ensure thoroughness and reproducibility (see Supplementary). For instance, in PubMed/MEDLINE, the search string was as follows: ('volleyball'[MeSH Terms] OR 'volleyball'[All Fields]) AND ('action quality'[All Fields] OR 'performance rating'[All Fields]) AND ('player position'[All Fields]).

Selection of sources of evidence

To facilitate a structured and transparent approach to screening and selection, all identified citations were imported into the CADIMA platform, an open-access, web-based tool specifically designed to support systematic reviews, and evidence maps across diverse disciplines.⁹ CADIMA was chosen due to its functionality in guiding reviewers through each stage of the scoping process, including duplicate removal, screening, documentation of inclusion decisions, and evidence tracking. Following the import, automatic de-duplication was conducted in CADIMA, and the remaining titles and abstracts were screened against the predefined eligibility criteria. A single author performed screening without secondary screening or consensus adjudication. As such, no procedures (e.g., dual screening, calibration exercises, or independent verification) were undertaken to reduce potential screening errors or reviewer bias.

After title and abstract screening, full texts of potentially relevant studies were retrieved and evaluated using CADIMA's built-in decision-logging interface. Again, the author solely conducted this step. Reasons for exclusion at the full-text stage were recorded in accordance with best practice standards outlined by PRISMA-ScR guidelines, though these decisions were not independently cross-checked. The limitations of this single-reviewer approach are acknowledged in the discussion section and are considered in the interpretation of results. Despite this limitation, CADIMA's structured environment enabled consistent documentation of decisions and preserved an auditable workflow. This transparency aligns with calls for improved methodological reporting in scoping reviews and enhances reproducibility.⁹

Data charting

Data charting was conducted using the CADIMA systematic review platform, which facilitated a structured, transparent, and reproducible approach to evidence extraction and synthesis. Given that this scoping review was conducted by a single researcher, no calibration exercise or inter-rater reliability procedures were implemented. However, to ensure methodological rigour, an iterative approach was maintained throughout the data charting process. The initial charting form was developed based on the

key elements outlined in the research objectives, including variables related to study metadata, rating system typologies, data collection methods, reliability and validity checks, and statistical approaches. As recommended by Kohl et al.,⁹ CADIMA's interface supported the systematic capture, storage, and revision of extracted information, enabling efficient tracking of modifications and rationales across the review timeline.

Throughout the charting process, the form was continuously refined to accommodate new insights and emerging categories from the included studies. This iterative revision was guided by both the methodological diversity of the literature and the need to capture nuanced features of action-quality rating systems in indoor volleyball. Specific attention was paid to ensuring consistency in variable definitions and the applicability of extracted information to the review's conceptual framework. All extracted data were reviewed multiple times during the synthesis stages to ensure completeness and coherence with the original sources.

Data items

To address the review's aim of mapping and categorising action-quality rating systems in indoor volleyball, a comprehensive set of data items was extracted from each included study. These items were selected to capture the methodological, contextual, and conceptual characteristics relevant to evaluating skill execution through performance analysis. The charting process focused on both descriptive metadata and domain-specific content to support both the mapping and synthesis phases.

First, general study characteristics were recorded, including the full article title, year of publication, primary institutional affiliation of the lead author, and the country where the study was conducted. This was followed by documentation of the study's stated aim or research purpose to understand the analytical focus and theoretical framing of the work. Population-related data were then extracted, specifying the gender of participants, their competitive level (e.g., elite, sub-elite, university, youth), and the specific sample characteristics, such as the number of matches, sets, rallies, or individual player actions analysed. Where applicable, the tournament or competition context was also recorded to situate the study within real-world match environments (e.g., national league matches, international tournaments, and Olympic Games).

Subsequently, for each study, the variables measured were identified and described, particularly those related to the assessment of skill quality. These included the specific skills evaluated (e.g., serve, pass, set, attack, block, dig), as well as any contextual variables (e.g., match phase, player position, score differential) used in the analysis. The methods of data collection were also documented, including the use of observational tools, video analysis

software, notational systems, and automated coding platforms. Furthermore, central to this review was the extraction of each study's action-quality rating system typology. The extracted typology included the scale structure (e.g., ordinal, interval, probabilistic), the number of categories or scoring levels, and the conceptual logic underlying the rating (e.g., outcome-based, decision-quality, biomechanical execution). Where available, we also noted whether the rating system was adapted from existing frameworks or developed independently.

Simultaneously, to assess methodological robustness, we recorded whether and how the studies addressed reliability and validity. Specifically, we noted any mention of inter- or intra-observer reliability measures, training procedures for coders, instrument validation steps, or statistical techniques used to establish the consistency or predictive accuracy of the rating system. We also captured the statistical approaches employed in the analysis, including descriptive, comparative, sequential, inferential, or modelling techniques (e.g., chi-square, ANOVA, regression, Markov chains, and Bayesian estimations). Finally, key findings and study implications were summarised, with particular attention to insights related to action-quality interpretation, rating system applicability, or relevance to coaching and match preparation. Where applicable, the format of data representation, such as tables, graphs, transition matrices, or coding trees, was recorded to understand how information was visualised and communicated. Therefore, these data items enabled a structured and nuanced mapping of how action quality has been conceptualised, operationalised, and applied within the volleyball performance analysis literature.

Critical appraisal of individual sources of evidence

Given that the objective of this scoping review is to identify and map the types of action-quality rating systems used to evaluate individual volleyball skills, a formal critical appraisal of the included studies was not conducted.^{6,7} The purpose of a scoping review is to provide a comprehensive overview of the available evidence, regardless of study quality, to examine the breadth, characteristics, and conceptual boundaries of the field. The present review, therefore, focused on describing and synthesising the approaches, typologies, and analytical frameworks employed across the literature rather than evaluating the methodological rigour of each source of evidence.

Synthesis of results

We conducted a structured narrative synthesis to collate and interpret findings across included studies, following established guidelines.^{10,11} This approach was selected due to the heterogeneity in rating system typologies, analytical techniques, and application contexts across studies. The synthesis aimed to answer four predefined research

questions related to the types, methods, contexts, and positional patterns of volleyball action-quality ratings. We first grouped studies according to rating system typology (i.e., ordinal-descriptive, interval-probabilistic, hybrid-emergent) to establish conceptual homogeneity. Within each group, we examined methodological features, including the statistical, computational, or inferential techniques used to interpret action-quality ratings. Analytical methods were then inductively categorised into four overarching domains: (a) descriptive and inferential statistics; (b) sequential and network-based analyses; (c) probabilistic and Bayesian modelling; and (d) machine learning and composite metric derivation. These categories were refined through iterative comparison with existing volleyball performance analysis frameworks.

Data were summarised in tabular format using a structured extraction form that included each study's analytical approach, primary variable targets, and interpretation of action-quality outcomes. Key themes and patterns were identified across studies through constant comparison, with attention paid to how methodological choices aligned with different rating purposes, game contexts, or positional foci. Where applicable, we highlighted methodological innovations, limitations, and implications for practice. We did not conduct a quantitative meta-analysis due to the non-comparability of outcome metrics and the exploratory aims of the review. This approach allowed us to (a) map the landscape of evaluative and analytical practices in volleyball action-quality research; (b) identify trends and gaps in methodological application; and (c) generate evidence-informed recommendations for future research and practitioner use.

Results

Selection of sources of evidence

A total of 305 records were identified through database searches across PubMed/MEDLINE ($n = 79$), Scopus ($n = 10$), and SPORTDiscus ($n = 216$). After the removal of 6 duplicate records, 299 titles and abstracts were screened; of these, 247 were excluded for not meeting the inclusion criteria. The remaining 52 full-text reports were retrieved and assessed for eligibility. Subsequently, 8 were excluded: 2 due to unavailability in English and 6 due to the under-representation of female participants. As a result, 44 studies fulfilled all the specified criteria and were included in the final scoping review (see Figure 1).

Characteristics of sources of evidence

The 44 included studies demonstrated broad international representation and variability in gender focus, competition context, and sample size. Geographically, the research was concentrated in eight countries. Spain accounted for

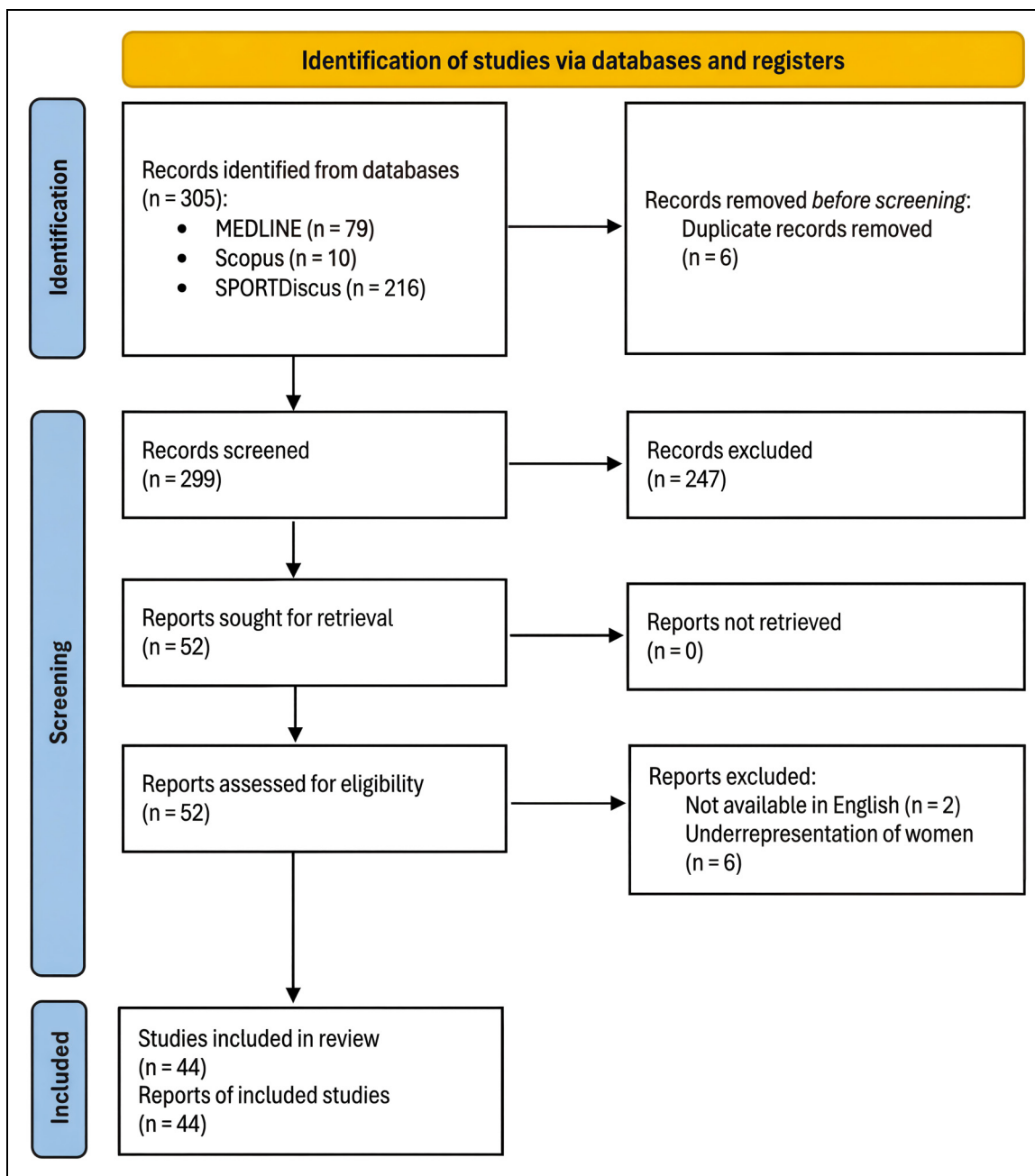


Figure 1. PRISMA 2020 flow diagram for the present review.

the largest proportion ($n = 15$; 34.1%),^{12–26} followed by the USA ($n = 8$; 18.2%),^{3,4,27–32} Brazil ($n = 7$; 15.9%),^{33–39} Greece ($n = 6$; 13.6%),^{40–45} Portugal ($n = 4$; 9.1%),^{46–49} Slovakia ($n = 2$; 4.5%),^{50,51} and one study each from Finland (2.3%)⁵² and Poland (2.3%).⁵³ Regarding the participants of each of the individual sources, most studies focused exclusively on female athletes ($n = 29$; 65.9%),^{3,4,12–15,17–20,24–26,29–32,34–39,43,46,47,49,50,52} with the remaining 15 including both men and women in either comparative or pooled analyses (34.1%; see Figure 2(a)).^{16,21–23,27,28,33,40–42,44,45,48,51,53}

Competition levels and tournament settings were categorised into five contextual groupings (see Figure 2(b)). International-level studies ($n = 22$; 50%)^{16–18,21–23,25,26,28,33,36,37,40,42–49,51} comprised analyses of the Olympic Games, FIVB World Championships, European Championships, and other top-tier events, providing high-performance benchmarks. National-level studies ($n = 10$; 22.7%)^{12,14,15,19,34,35,38,39,41,53} examined domestic elite leagues such as the Brazilian Superliga and Italian Serie A1. Collegiate-level studies ($n = 7$; 15.9%)^{3,4,27,29–32} focused primarily on NCAA Division I programs in the USA. A

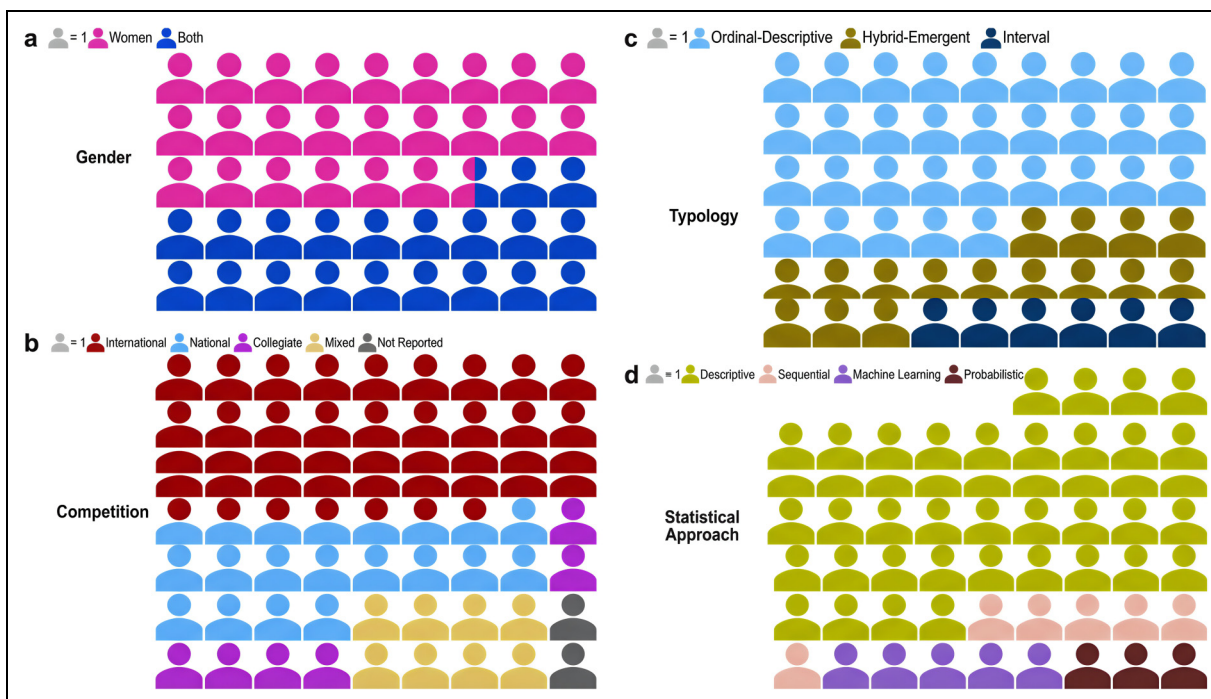


Figure 2. Number of studies by their characteristics included in the present review.

small subset ($n = 4$; 9.1%)^{13,20,50,52} combined elite and junior categories within the same design, enabling cross-level comparisons. One study did not specify its competition context (2.3%).²⁴

Sample sizes and units of analysis varied across studies. Some focused on large datasets encompassing hundreds of matches and thousands of skill executions (e.g., 100+ matches).^{4,12,20,27,28,34,39,43,44,48,53} In contrast, others employed smaller, focused samples (e.g., 6–8 matches)^{15,24–26,46,47,52} to explore specific tactical or positional questions. Further details about the individual studies, including specific data on gender focus, competition contexts, and sample sizes, are provided in Supplementary. Therefore, these characteristics emphasise the heterogeneity and richness of the evidence base, spanning multiple countries, genders, levels of play, and analytical scopes. This diversity enhances the transferability of insights while also accentuating contextual nuances that may influence action-quality measurement in volleyball.

Results of individual sources of evidence

The included studies employed a diverse array of rating typologies, analytical frameworks, and positional emphases in evaluating action quality in elite-level volleyball. This section presents the findings across three domains: (a) rating system typologies, (b) analytical approaches, and (c) action quality analysis by player position.

Rating system typologies. Most studies ($n = 28$; 63.6%; see Figure 2(c))^{13,14,16,19,20,22–26,31–35,38,40–46,48,50–53} employed ordinal-descriptive rating systems grounded in established coding scales, such as those developed by the FIVB.^{1,54–57} These systems assigned fixed values to actions (e.g., 1–4 or 1–5 scale) based on subjective or operationalised definitions of quality, often using discrete, skill-specific metrics (e.g., serve efficacy, dig rating) to evaluate performance across contexts. Subsequently, a smaller subset ($n = 4$; 9.1%)^{4,28–30} applied interval-scaled or continuous rating models, particularly when modelling skill transitions or action outcomes probabilistically. These studies moved beyond ordinal scales by applying continuous probability estimates to rating categories, often using posterior distributions or logistic coefficients as interpretive metrics. Twelve studies (27.3%)^{3,12,15,17,18,21,27,36,37,39,47,49} developed or adopted hybrid frameworks combining descriptive metrics with context-specific weighting, social network analysis, or emergent coding logic, using interaction networks, eigenvector centrality, and sequential-action typologies to assess game-phase efficacy. Further details about the individual studies, including instruments used and variables measured, are provided in Supplementary.

Analytical approaches. The majority of studies employed traditional descriptive statistics ($n = 26$; 59.1%; see Figure 2(d))^{13,14,16,17,19,20,22–25,31–35,38,41–45,48,50–53} followed by inferential tests such as t-tests^{2,2,44,45} chi-square tests,^{13,22,23,25,31,32,50} and ANOVAs.^{19,34,38,41,53}

Regression-based approaches^{17,20,33,41,42} were frequently used to identify predictors of performance outcomes. A growing body of work ($n = 8$; 18.1%)^{15,36,37,39,40,46,47,49} integrated temporal and relational dynamics via sequence and network analysis. Sequential lag analysis,⁴⁶ game-complex sequencing,^{15,39,47} and centrality measures^{36,37,49} were used to uncover patterns of action transitions and the structural properties of rallies. Several studies ($n = 5$; 11.4%)^{3,4,28–30} applied probabilistic and Bayesian models to capture uncertainty and transition dynamics. Recent studies ($n = 5$; 11.4%)^{12,18,21,26,27} integrated machine learning to derive composite efficiency metrics.^{18,27} These works utilised recursive partitioning,²¹ kernel density estimation,²⁷ and unsupervised clustering¹² to build composite indices. Further information about the individual sources, including reliability and validity measures, are provided in Supplementary.

Action-Quality Ratings Across Player Positions. Several studies examined how action-quality ratings varied by playing position, providing position-specific insights into technical demands and performance contributions. Key comparisons show that setters were evaluated for more than technical accuracy, with Hybrid and probabilistic frameworks incorporating spatial variability, decision-making complexity, and downstream influence on attack outcomes.^{16,21,39} In contrast, outside and opposite hitters were most frequently represented, with analyses emphasising their action volume, attack versatility across zones and tempos, and terminal efficacy.^{42,53} Outside hitters, specifically, were more involved in transition and coverage actions, while studies also reported role-based variations in serve type, serving zone, and effectiveness.^{25,26} Middle blockers were typically analysed for block participation, tempo-dependent attack execution, and positional decision-making within defensive systems.^{13,38} Defensive specialists like liberos, though rarely the focus of full performance models, featured prominently in continuity-based metrics, including dig

performance, phase-specific reliability, and reception quality in different match contexts.^{14,18,39} Notably, liberos had higher pass accuracy in early set or lower-pressure situations, but performance dropped under score-line pressure.^{14,18,42} Despite the rise of Hybrid-Emergent rating systems, literature still focuses primarily on setters and attackers, with liberos and middle blockers comparatively underrepresented in multi-action Hybrid frameworks, despite their tactical importance. Detailed results from individual sources are in Supplementary Table B4.

Evolution impact on coaching and analysis practices. The progression from ordinal-descriptive to probabilistic and hybrid frameworks represents more than methodological advancement (see Figure 3), it reflects a fundamental shift in how volleyball performance is conceptualised and applied:

1. **Strategic Decision-Making:** Early ordinal systems (e.g., FIVB 1–4 scales) provided binary success/failure feedback, limiting strategic insights. Modern probabilistic approaches (e.g., Markov chain models used in 11.4% of studies) enable coaches to quantify transition probabilities and optimise tactical decisions based on expected point values rather than subjective assessments.
2. **Player Development:** The shift toward contextual sensitivity allows coaches to identify specific situational strengths and weaknesses. For instance, studies using network-based analysis revealed that action quality propagates differently through rally phases, enabling targeted intervention strategies for phase-specific skill deficits.
3. **Match Preparation:** Advanced composite metrics facilitate opponent analysis by revealing systematic patterns in action quality under varying match conditions, pressure situations, and positional rotations.

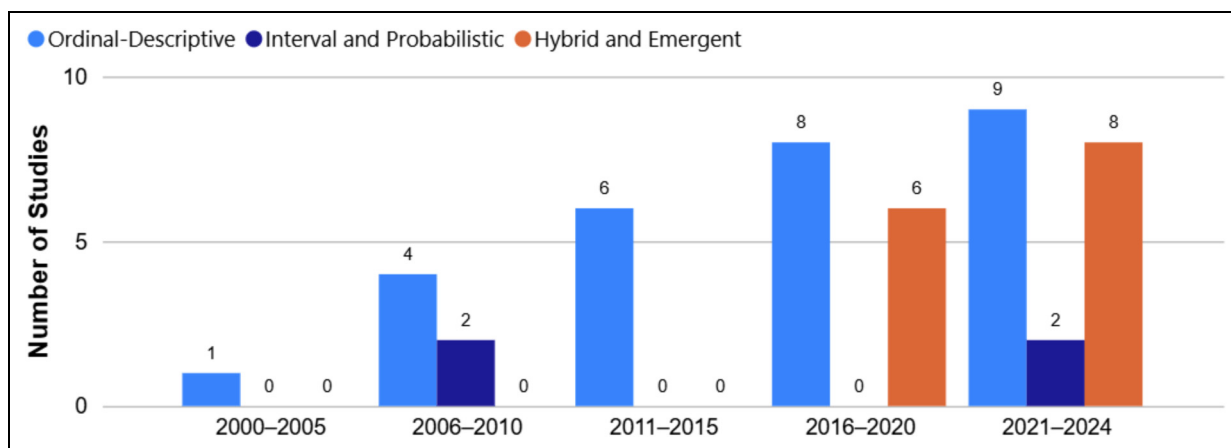


Figure 3. Temporal distribution of action-quality rating system typologies (2000–2024).

Synthesis of results

The analytical interpretation of action-quality ratings in volleyball has evolved from frequency-based, inferential statistics to context-sensitive, probabilistic, and machine-learning approaches. Each methodological class contributes uniquely to understanding action quality: (a) descriptive-inferential methods offer interpretability and ease of application but may overlook temporal complexity; (b) sequential and network approaches enhance understanding of interaction structure across game phases; (c) probabilistic and Bayesian models frame quality in terms of expected impact and uncertainty; and (d) composite and machine-learning techniques operationalise multidimensional performance models suited to modern analytics. The synthesis suggests that no single method is sufficient; rather, integrative approaches that combine statistical rigour with structural and contextual sensitivity are best suited for contemporary action-quality evaluation in volleyball.

Discussion

Summary of evidence

Typologies of rating systems. The dominant typological foundation across the 44 studies was ordinal-descriptive rating, used in over 60% of included research. These systems apply discrete, rank-ordered categories to skill execution (e.g., 'poor,' 'adequate,' 'excellent'). They are commonly adapted from legacy frameworks such as the FIVB's statistical system^{1,54-56} and Eom and Schultz's⁵⁷ evaluative scale. These ordinal systems often use 4-6 point Likert-type scales, with skill-specific operational definitions (e.g., serve efficacy, attack continuity) delineated by outcome or technique quality. While early studies applied uniform scales across skills, later work adopted skill-specific categorical rubrics, allowing finer discrimination for actions like setting or digging, particularly in complex transitions or out-of-system play. Additionally, emergent typologies included interval and probabilistic models, introduced to address the limitations of fixed ordinal categories in capturing within-category variation and interaction dynamics. These systems, most notably Markov chains,^{4,29,30} Bayesian regressions,^{3,4,28} and expected point value models, framed each technical action as part of a probabilistic decision tree or event sequence. Here, rating is not an evaluative label but an estimated transition likelihood or point differential derived from observed frequencies and modelled priors. Furthermore, a third class, hybrid and emergent frameworks, integrated spatial-temporal, contextual, and positional inputs to derive composite indices such as the Technical Individual Contribution Coefficient (T-ICC),¹⁸ Points Scored Above Average per Set (PAAPS),²⁸ and context-indexed weighted performance scores. These systems treated performance as an emergent property of interaction networks, often informed by game-complex sequencing, phase-of-play logic (K0-KV),

and spatial positioning. This typology reflects an epistemological shift, from viewing action quality as static and isolated to seeing it as dynamic, probabilistic, and contingent on match phase, role, and task constraints.

Analytical approaches. The analytical techniques employed across studies mapped closely to the complexity of their rating systems. Descriptive statistics (means, standard deviations, frequencies) were ubiquitous, especially in ordinal-descriptive studies where they served to summarise action frequencies and skill efficacy distributions. Basic inferential techniques (i.e., t-tests, Chi-square, ANOVA) were frequently used to compare outcomes across genders, competition levels, or match contexts. However, their interpretive scope was often constrained by assumptions of variable independence and scale linearity. Moreover, to address inter-variable dependency and multivariate dynamics, a growing subset of studies employed multivariate techniques such as MANOVA, canonical discriminant analysis,⁴⁸ and multinomial logistic regression.^{33,35} These enabled researchers to identify combinations of action qualities that best predicted performance outcomes (e.g., side-out success, attack efficacy), and were especially useful in classifying team or positional profiles. Subsequently, more advanced studies applied Bayesian and machine learning methods to model uncertainty and optimise classification. For example, Florence et al.²⁹ and Miskin et al.³⁰ used Dirichlet priors within Markov-chain frameworks to estimate transition probabilities with quantified uncertainty, while Fellingham²⁸ used hierarchical logistic models to derive positional contributions to rally success. Several studies used Monte Carlo simulations, Gibbs sampling, or recursive partitioning to test the stability and predictive accuracy of their models, a methodological advancement rarely seen in earlier performance-analysis literature. Simultaneously, an emerging approach was the use of network-based analysis,^{36,47} where game actions were treated as nodes and transitions as edges, enabling calculation of centrality measures (e.g., eigenvector, degree centrality) to identify key skills and phases within game-complex architectures. This allowed for visualisation and quantification of how action quality propagates through a rally, offering unique insights into the systemic structure of elite play.

Implementation contexts. The application of action-quality rating systems spanned a wide array of competitive, demographic, and situational contexts, providing insight into how rating validity may vary across volleyball subpopulations. Most studies examined elite-level women's volleyball, particularly within international competitions (e.g., FIVB World Championships, Olympics, continental leagues), where data infrastructure, match filming, and notational access were most developed. These contexts often allowed for richer, multi-angle video analysis and greater sample sizes, enhancing the precision and reliability of coding procedures. Conversely, several studies^{24,50} investigated youth, junior, or developmental categories, including under-18, collegiate, or sub-elite national players. These

studies underlined different action-quality profiles, often lower error rates but less tactical variation, and suggested that rating scales must account for developmental constraints and learning curves. Furthermore, match-specific variables such as set number, scoreline, game phase, and match importance (e.g., finals vs. group stage) were increasingly embedded within more recent studies.^{18,42} This reflects a movement toward phase-sensitive and context-aware performance indicators, in contrast to earlier models that assumed rating consistency across match segments. These contextualised models improve ecological validity, enabling inferences about action quality under fatigue, pressure, or shifting tactical demands. Geographically, the sample drew heavily from European, South American, and North American volleyball cultures, with Spain, Portugal, Brazil, and the United States frequently represented. Most studies were situated within academic–federation partnerships, granting access to national databases and league systems, though transparency in data access and sample origin varied across studies.

Positional and skill-specific focus. Analysis of action-quality ratings was strongly skewed toward serve, reception, set, attack, dig, and block, with terminal actions (e.g., spike, serve) most frequently evaluated. However, skill evaluations were often disaggregated by player position, illuminating unique tactical demands and rating requirements. First, setters were typically assessed on setting location, accuracy, attack tempo enabled, and response to defensive disruption.^{36,38} Metrics such as setter movement distance, setter's action range with availability of first tempo (SARA),²¹ and the ability to maintain first-tempo attack under suboptimal conditions were frequently used. Advanced analyses included conditional dependencies between dig quality and setting zone,⁴⁵ or decision-tree modelling of side-out options by setter rotation. Second, outside hitters and opposites were central to terminal efficiency evaluations. Studies such as Drikos et al.⁴² and Palao et al.²² emphasised attack outcome variation by position, with outside hitters operating more often under pressure (e.g., broken play, off-system sets) and opposites demonstrating high-risk serving and back-row attack efficacy. These roles were often paired with block metrics, reflecting their dual offensive and defensive impact. Last, liberos and defensive specialists were typically assessed on reception quality, dig location, and pass-to-attack transition facilitation. Ratings often used categorical scales with binary error scoring. However, newer frameworks (e.g., T-ICC R3)¹⁸ began weighting their contribution to systemic outcomes, such as enabling middle attack or shifting setter location. Notably, libero actions were underrepresented in early ordinal studies but gained analytical weight in emergent frameworks focused on first-ball continuity and out-of-system resilience. Hence, the positional granularity of action-quality ratings is increasing, moving from generic

skill execution toward role-sensitive metrics that reflect systemic contribution and tactical function. However, few studies have yet unified these insights into fully integrated, team-wide efficiency frameworks, marking a key direction for future work.

Limitations

Several methodological and interpretive limitations of this scoping review warrant consideration. The principal limitation is that screening, selection, and data extraction were performed by a single reviewer, without independent duplicate screening, consensus adjudication, or retrospective verification. Although this approach is sometimes adopted in resource-constrained scoping reviews and was transparently reported, it may increase the risk of selection bias. Subjectivity may have influenced decisions during title/abstract screening, full-text eligibility assessment, and data charting, particularly in cases requiring interpretive judgment rather than clear-cut decisions.

This limitation is especially salient with respect to the classification of action-quality rating system typologies, and most critically, the identification of Hybrid-Emergent frameworks. Unlike ordinal or probabilistic systems, Hybrid-Emergent typologies are defined by conceptual blending, adaptive weighting, or post-hoc analytical evolution, characteristics that are often implicitly described rather than formally labelled in primary studies. As a result, typological assignment required interpretive synthesis across methodological descriptions, analytical procedures, and reported use-cases. Without secondary reviewer verification or inter-rater reliability assessment, there remains the possibility that some systems were misclassified, over-interpreted, or inconsistently grouped, which could influence the apparent prevalence and conceptual boundaries of this typology.

The absence of duplicate data extraction prevented calculation of inter-rater agreement, such as Cohen's κ , and limits confidence in whether extracted variables – rating scale structure, reliability reporting, and analytical application – are reproducible. Although structured charting and clear operational definitions reduced error, these do not fully substitute for independent checks. Therefore, the synthesis is best seen as a theoretically informed overview, not a definitive taxonomy of volleyball action-quality rating systems.

Additional limitations reflect the scope and intent of scoping reviews. Consistent with scoping review methodology, no formal appraisal of methodological quality or risk of bias was undertaken; however, this does hinder evaluation of internal validity or weighting of findings. The search strategy was limited to peer-reviewed, English-language literature, thereby excluding grey literature, such as coaching manuals, federation reports, and applied analytics frameworks. Given the practice-led nature

of performance analysis research, this exclusion may lead to insufficient representation of innovative or field-embedded rating systems not available in academic publication avenues.

Considerable heterogeneity across studies – in rating system design, analytical methods, competitive context, and population – necessitated a narrative synthesis. Although findings were systematically organised by typology, analytical logic, and contextual application, such diversity constrains comparability and precludes both quantitative aggregation and generalisability. This review does not evaluate predictive validity or performance impact, nor does it advocate frameworks. Principal limitations, therefore, include lack of comparability, restricted generalisability, and the absence of predictive performance assessment.

Future research should prioritise dual-reviewer screening and extraction, formal assessment of inter-rater agreement, and structured expert consultation, as current approaches may be limited by single-reviewer bias and subjective boundary definitions. These steps would help refine typological boundaries, particularly for emergent hybrid systems. Integrating methodological appraisal and outcome-based validation would further strengthen the evidence base, addressing the limitation of inconsistent appraisal standards. Doing so would also support translation into high-performance coaching and talent development environments while mitigating risks of misapplication due to insufficient validation.

Practical applications for coaches and analysts

The findings of this review provide several actionable insights for volleyball practitioners:

Performance evaluation framework selection. Based on our analysis, coaches should select rating systems that align with their specific objectives. For positional skill assessment, ordinal-descriptive systems (used in 63.6% of studies) offer interpretability and ease of implementation. However, for comprehensive team analysis, hybrid frameworks that integrate contextual variables provide more nuanced insights into player contributions within game systems.

Training program design. The evolution from frequency-based to context-sensitive approaches suggests coaches should incorporate situational variables when assessing skill execution. For example, setter evaluation should consider not only technical accuracy but also spatial variability and attack outcome facilitation, as demonstrated in advanced frameworks like the Technical Individual Contribution Coefficient.

Talent identification enhancement. The position-specific rating patterns identified in this review (e.g., liberos showing higher pass accuracy in low-pressure situations) can inform recruitment strategies and position-specific development pathways. Coaches can use these benchmarks to identify players who excel under specific match conditions

Conclusion

This scoping review offers a comprehensive map of how action-quality rating systems have been conceptualised, operationalised, and applied in volleyball performance analysis. Drawing upon 44 included studies, our synthesis revealed that the field has evolved from foundational ordinal-descriptive frameworks to more nuanced, probabilistic, and hybrid systems that integrate technical actions with contextual and systemic match dynamics. These developments mark a conceptual adjustment: from static skill categorisation toward dynamic modelling of player actions within evolving team systems and match contexts. Subsequently, a primary insight is the increasing use of emergent composite metrics, which reflect a more comprehensive approach to evaluating action quality. These models not only grade execution quality but embed the skill within its functional role, such as whether a high-quality set facilitates a first-ball side-out or contributes to phase transition effectiveness. Hence, the emergence of such composite, context-sensitive metrics reflects a paradigmatic shift in volleyball analytics, from isolated skill assessment to system-sensitive performance modelling. From an analytical standpoint, we observed an expansion in methodological sophistication. While many early studies employed descriptive statistics and ordinal coding systems, recent research increasingly utilises machine learning, multivariate regressions, Markov chains, and dynamic network analyses. This transition towards diverse analytical approaches reflects a methodological pluralism responsive to the complexity of performance in elite volleyball. Additionally, several studies linked rating systems to performance indicators such as rally outcomes, phase success, or role-specific contributions, thereby increasing their applied utility for coaches and analysts.

Accordingly, implications for practice are multifaceted. First, action-quality rating systems must be designed with contextual and positional specificity. Studies that evaluated setters, liberos, or opposites in isolation frequently reported that positional responsibilities mediate the meaning and impact of technical execution. As such, rating frameworks that fail to accommodate role-differentiated performance risk producing erroneous conclusions. Second, our findings suggest that future system development should balance interpretability with complexity. While machine-learned models offer strong predictive power, their lack of transparency may limit adoption in applied settings. A hybrid

approach (i.e., leveraging both data-driven and theory-informed metrics) may provide a more useful bridge between research innovation and coaching practice. Finally, this review underlines significant gaps in the literature, particularly regarding validation across populations (e.g., youth vs. elite, male vs. female), under-representation of defensive and transition phases, and inconsistent terminology across studies. Thus, addressing these gaps is vital for the next generation of rating systems to serve both academic and applied purposes.

In conclusion, the trajectory of volleyball action-quality rating research reflects growing methodological depth and increasing alignment with systemic performance models. However, meaningful advancement will require standardisation of terminology, robust validation across contexts, and deeper engagement with coaches and practitioners to ensure analytic outputs are both insightful and actionable. This scoping review lays the foundation for those future developments by consolidating what is known, revealing what is missing, and suggesting how the field might evolve.

ORCID iD

Kirsten N Spencer  <https://orcid.org/0000-0002-3188-2179>

Ethical considerations

There are no human participants in this article and informed consent is not required.

Consent to participate

There are no human participants in this article and informed consent is not required.

Consent for publication

Not applicable.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding statement

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Data availability

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Supplemental material

Supplemental material for this article is available online.

References

1. Coleman JE. *A statistical evaluation of selected volleyball techniques at the 1974 World's Volleyball Championships*. PhD Thesis, Brigham Young University, USA, 1975.
2. Palao JM, Manzanares P and Ortega E. Design, validation, and reliability of an observation instrument for technical and tactical actions in indoor volleyball. *Eur J Hum Mov* 2015; 34: 75–95.
3. Hass Z and Craig BA. Exploring the potential of the plus/minus in NCAA women's volleyball via the recovery of court presence information. *J Sports Anal* 2018; 4: 285–295.
4. Powers S, Stancil L and Consiglio N. Estimating individual contributions to team success in women's college volleyball. *J Quant Anal Sports* 2024; 21: 117–135.
5. Silva M, Marcelino R, Lacerda D, et al. Match analysis in volleyball: a systematic review. *Monten J Sports Sci Med* 2016; 5: 35–46.
6. Peters MDJ, Godfrey C, McInerney P, et al. Scoping reviews. In: E Aromataris, C Lockwood, K Porritt, et al. (eds) *JBI Manual for evidence synthesis*. Adelaide, Australia: JBI, 2020, pp.417–474.
7. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473.
8. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016; 75: 40–46.
9. Kohl C, McIntosh EJ, Unger S, et al. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: A case study on CADIMA and review of existing tools. *Environ Evid* 2018; 7: 9–53.
10. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *Br Med J* 2020; 368: 16890.
11. Popay J, Roberts H, Sowden A, et al. Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme. *J Epidemiol Community Health* 2006; 59. DOI: 10.1136/2005.011111
12. Casimiro-Artés MÁ, Hileno R and García-de-Alcaraz A. Applying unsupervised machine learning models to identify serve performance related indicators in women's volleyball. *Res Q Exerc Sport* 2024; 95: 47–53.
13. Echeverría C, Ortega E and Palao JM. Normative profile of the efficacy and way of execution for the block in women's volleyball from under-14 to elite levels. *Monten J Sports Sci Med* 2020; 9: 41–47.
14. García-de-Alcaraz A and Usero L. Influence of contextual variables on performance of the libero player in top-level women's volleyball. *J Hum Kinet* 2019; 70: 199–207.
15. Hileno R, Arasanz M and García-de-Alcaraz A. The sequencing of game complexes in women's volleyball. *Front Psychol* 2020; 11. DOI: 10.3389/fpsyg.2020.00739
16. López E, Rodrigo MV and Gea-García GM. The setter's attack in high-level volleyball. *Retos* 2024; 56: 95–106.

17. López-Serrano C, Zakythinaki M, Mon-López D, et al. Contextualizing evaluation of performance in volleyball: introducing contextual individual contribution coefficients to assess technical actions. *Percept Mot Skills* 2023; 130: 2663–2684.
18. López-Serrano C, Zakythinaki M, Mon-López D, et al. Introducing the technical individual contribution coefficient: a metric for evaluating performance in elite volleyball. *Int J Perform Anal Sport* 2024; 24: 204–217.
19. Campo V L-D, Cabeza E and Damas JS. Knowledge of opposition attack tendencies on volleyball blocking. *Rev Int Med Cienc Act Fis Deporte* 2022; 22: 845–861.
20. Marzano-Felisatti JM, Quesada JIP and Luján JFG. Women's volleyball performance indicators according to age category and team's final position in international competitions. *Eur J Hum Mov* 2022; 48: 4–20.
21. Millán-Sánchez A, Parra-Royón MJ, Benítez JM, et al. Ability to predict side-out performance by the setter's action range with first tempo availability in top European male and female teams. *Int J Environ Res Public Health* 2020; 17: 6326.
22. Palao JM, Ahrabi-Fard SJ and Ureña A. Effect of team level on skill performance in volleyball. *Int J Perform Anal Sport* 2004; 4: 50–60.
23. Palao JM, Manzanares P and Ortega E. Techniques used and efficacy of volleyball skills in relation to gender. *Int J Perform Anal Sport* 2009; 9: 281–293.
24. Palao JM, Ureña A, Moreno MP, et al. Effect of changes in the net height, court size, and serve limitations on technical-tactical, physical, and psychological aspects of U-14 female volleyball matches. *Front Psychol* 2024; 14. DOI: 10.3389/fpsyg.2023.1341297
25. Quiroga ME, García-Manso JM, Rodríguez-Ruiz D, et al. Relation between in-game role and service characteristics in elite women's volleyball. *J Strength Cond Res* 2010; 24: 2316–2321.
26. Quiroga ME, Rodríguez-Ruiz D, Sarmiento S, et al. Characterisation of the main playing variables affecting the service in high-level women's volleyball. *J Quant Anal Sports* 2012; 8. DOI: 10.1515/1559-0410.1348
27. Bagley C and Ware B. Bump, set, spike: Using analytics to rate volleyball teams and players. In: MIT Sloan Sports Analytics Conference, Hynes Convention Centre, Boston, MA, 2017.
28. Fellingham GW. Evaluating the performance of elite level volleyball players. *J Quant Anal Sports* 2022; 18: 15–34.
29. Florence LW, Fellingham GW, Vehrs PR, et al. Skill evaluation in women's volleyball. *J Quant Anal Sports* 2008; 4. DOI: 10.2202/1559-0410.1105
30. Miskin MA, Fellingham GW and Florence LW. Skill importance in women's volleyball. *J Quant Anal Sports* 2010; 6. DOI: 10.2202/1559-0410.1234
31. Palao JM. Side-out success and ways that points are obtained in women's college volleyball. *J Sports Anal* 2018; 4: 243–250.
32. Palao JM and Ahrabi-Fard I. Side-out success in relation to setter's position on court in women's college volleyball. *Int J Appl Sports Sci* 2011; 23: 155–167.
33. Araújo CRP, Tosini L, Freire AB, et al. Reception-attack relation in men's and women's volleyball during the Rio 2016 Olympics. *J Phys Educ Sport* 2020; 20: 2008–2012.
34. Campos FAD, Stanganelli LCR, Pasquarelli BN, et al. Performance indicators analysis at Brazilian and Italian women's volleyball leagues according to game location, game outcome, and set number. *Percept Mot Skills* 2014; 118: 347–361.
35. Costa GDCT, Afonso J, Barbosa RV, et al. Predictors of attack efficacy and attack type in high-level Brazilian women's volleyball. *Kinesiology* 2014; 46: 242–248.
36. Laporta L, Medeiros AIA, Vargas N, et al. Coexistence of distinct performance models in high-level women's volleyball. *J Hum Kinet* 2021; 78: 161–173.
37. Laporta L, Valongo B, Afonso J, et al. Game-centred study using eigenvector centrality in high-level women's volleyball: play efficacy is independent of game patterns... or is it? *Monten J Sports Sci Med* 2021; 10: 19–24.
38. Rocha ACR, Pedrosa GF, Freire AB, et al. Analysis of the setting and predictive factors of the effect of attack according to game ecology: the case of female volleyball. *Kinesiology* 2020; 52: 217–223.
39. Vieira PHC, Laporta L, Pedrosa GF, et al. Sequencing of game complexes in women's volleyball: are there differences according to the set played? *Hum Mov* 2024; 25: 96–104.
40. Barzouka K, Sotiropoulos K, Drikos S, et al. Current trends of the serve skill in relation to the in-game roles of the elite volleyball players: comparison between genders. *J Hum Sport Exerc* 2021; 16: 317–331.
41. Drikos S, Sotiropoulos K, Barzouka K, et al. The contribution of skills in the interpretation of a volleyball set result with minimum score difference across genders. *Int J Sports Sci Coach* 2020; 15: 542–551.
42. Drikos S, Sotiropoulos K, Gkreka S, et al. Variations in attack patterns between female and male outside hitters in top-level volleyball. *Int J Sports Sci Coach* 2023; 18: 245–256.
43. Giatsis G. Performance indicators in women's volleyball Olympics and World Championships (2014–2021). *Int J Sports Sci Coach* 2023; 18: 1266–1276.
44. Kountouris P, Drikos S, Aggelonidis I, et al. Evidence for differences in men's and women's volleyball games based on skills effectiveness in four consecutive Olympic tournaments. *Compr Psychol* 2015; 4. DOI: 10.2466/30.50.Cp.4.9
45. Sotiropoulos K, Barzouka K, Tsavdaroglou S, et al. Comparison and assessment of the setting zone choices by elite male and female volleyball setters in relation to the quality of defence. *Facta Univ Ser Phys Educ Sport* 2019; 17: 57–68.
46. Afonso J, Mesquita I, Marcelino R, et al. Analysis of the setter's tactical action in high-performance women's volleyball. *Kinesiology* 2010; 42: 82–89.
47. Hurst M, Loureiro M, Valongo B, et al. Systemic mapping of high-level women's volleyball using social network analysis: the case of serve (K0), side-out (KI), side-out transition (KII) and transition (KIII). *Int J Perform Anal Sport* 2016; 16: 695–710.

48. João PV, Leite N, Mesquita I, et al. Sex differences in discriminative power of volleyball game-related statistics. *Percept Mot Skills* 2010; 111: 893–900.
49. Laporta L, Afonso J and Mesquita I. Interaction network analysis of the six game complexes in high-level volleyball through the use of eigenvector centrality. *PLoS One* 2018; 13. DOI: 10.1371/journal.pone.0203348
50. Adamec P and Přidal V. The relationship between serve-receive quality and attack performance across different age categories in women's elite volleyball. *J Phys Educ Sport* 2024; 24: 1133–1138.
51. Přidal V and Priklerová S. Analysis of relation between team placing in tournament and selected indicators of playing performance in top-level volleyball. *J Phys Educ Sport* 2018; 18: 1501–1505.
52. Inkinen V, Häyrinen M and Linnamo V. Technical and tactical analysis of women's volleyball. *Biomed Hum Kinet* 2013; 5: 43–50.
53. Pawlik D, Bobula G and Mroczek D. The effectiveness and types of serves used in elite women's and men's volleyball in the 2021/2022 season. *Sci Rep* 2024; 14. DOI: 10.1038/s41598-024-68262-5
54. Coleman JE, Neville B and Gordon B. A statistical system for volleyball and its use in Chicago Women's Association. *Int Volleyball Rev* 1969; 17: 72–73.
55. Díaz J. *Voleibol: La dirección de equipo [Volleyball: Team management]*. 2nd ed. Sevilla, Spain: Wanceulen Editorial S.L., 1992.
56. Díaz J. *Análisis y significación de los comportamientos técnicos, tácticos y competitivos de voleibol masculino en los juegos de la XXV olimpiada de Barcelona [Analysis and significance of technical, tactical, and competitive behaviors in men's volleyball at the XXV Olympic Games in Barcelona]*. PhD Thesis, University of Seville, Spain, 1996.
57. Eom HJ and Schutz RW. Statistical analyses of volleyball team performance. *Res Q Exerc Sport* 1992; 63: 11–18.