

Statistical modelling in exercise and sport sciences: methodological challenges, decision-oriented frameworks, and applied solutions

review paper

DOI: <https://doi.org/10.5114/hm/225606>

© Wrocław University of Health and Sport Sciences

GIANPIERO GRECO 

Department of Translational Biomedicine and Neuroscience (DiBrain), University of Bari Aldo Moro, Italy

ABSTRACT

Exercise and sport sciences increasingly rely on statistical models to analyse complex data arising from experimental, observational, and longitudinal designs. However, applied research frequently shows misalignment between research questions, data structure, and analytical strategies, leading to biased inference, limited reproducibility, and reduced interpretability. Common challenges include small sample sizes, clustered and repeated measurements, suboptimal model selection, insufficient assessment of model assumptions, and the growing use of machine learning methods without transparent validation or reporting. This methodological narrative review synthesises key statistical challenges through a targeted literature search and proposes a decision-oriented framework to guide statistical model selection and reporting in exercise and sport sciences. The framework integrates study objective (explanatory vs. predictive), outcome type, data structure, and information content to support coherent and transparent analytical decisions. Classical models, penalised regression, and supervised machine learning approaches are discussed in terms of applicability, assumptions, interpretability, and validation requirements. A workflow figure and practical decision tables are provided to support model selection, test selection, validation, and robustness assessment. The overarching goal is to strengthen applied statistical reasoning and promote robust, interpretable, and reproducible quantitative research in exercise and sport sciences.

Key words: statistical inference, mixed-effects models, machine learning, validation, reproducibility

Introduction

Exercise and sport sciences increasingly rely on quantitative methods to investigate complex relationships between physical activity, health outcomes, and human performance. Contemporary research designs frequently involve repeated measurements, clustered observations (e.g., teams, schools, training groups), multidimensional outcomes, and heterogeneous populations. These characteristics reflect the inherently multi-level and dynamic nature of exercise-related phenomena, in which individual responses are embedded within contextual, temporal, and organisational structures. As a consequence, statistical analyses must explicitly account for dependency, heterogeneity, and temporal dynamics to avoid biased estimates and misleading conclusions. While advances in statistical modelling and data-driven techniques have expanded the analytical toolbox available to researchers, their application in exercise and sport sciences remains methodologi-

cally inconsistent and, in many cases, insufficiently transparent.

Recent methodological discussions in exercise and sport sciences have highlighted persistent concerns related to transparency, reproducibility, statistical power, reporting practices, questionable inferential approaches, and the need for stronger collaboration between applied researchers and statisticians [1–6]. These issues include inadequate alignment between research questions and statistical models, insufficient sample-size planning, inappropriate handling of correlated or repeated data, limited assessment of model assumptions, and overreliance on dichotomous or insufficiently justified interpretations of statistical evidence. Such limitations are particularly critical in applied contexts, where statistical results are often used to inform training prescriptions, clinical decision-making, rehabilitation protocols, and public health recommendations. In these settings, weak statistical practices may translate directly into suboptimal or ineffective interven-

Correspondence address: Gianpiero Greco, Department of Translational Biomedicine and Neuroscience (DiBrain), University of Bari Aldo Moro, 70124 Bari, Italy, e-mail: gianpiero.greco@uniba.it; <https://orcid.org/0000-0002-5023-3721>

Received: February 02, 2026

Accepted: July 08, 2026

Citation: Greco G. Statistical modelling in exercise and sport sciences: methodological challenges, decision-oriented frameworks, and applied solutions. Hum Mov. 2026;27(4):1–16; doi: <https://doi.org/10.5114/hm/225606>.

tions. These challenges are further amplified by the increasing adoption of machine learning approaches, which, despite their predictive potential, are often implemented as “black-box” models with limited interpretability and insufficient validation [7, 8].

In this context, methodological rigour in exercise and sport sciences depends not solely on the sophistication of statistical techniques, but on the coherence between study objectives, data structure, and analytical decisions. A mismatch between these elements may lead to biased inference, inflated performance estimates, and reduced reproducibility across studies. Recent methodological literature emphasises the need to distinguish clearly between explanatory and predictive modelling, to account explicitly for dependency structures in the data, and to adopt transparent and reproducible statistical practices [9, 10].

Contemporary methodological contributions highlight the need to move beyond dichotomous interpretations of statistical evidence and instead promote estimation-focused, decision-relevant inference. This shift is particularly important in applied health and sport science research, where nuanced interpretation of uncertainty is essential for informing practice. McShane *et al.* [11] argue that the conventional threshold-based approach to statistical significance obscures practical relevance and encourages misleading binary conclusions, advocating instead for an integrative evidential perspective in which p -values are interpreted alongside effect sizes, confidence intervals, direction and magnitude of effects, model assumptions, sensitivity analyses, practical relevance, and consistency with prior evidence. In a sport science context, this means evaluating whether an observed effect is not only statistically detectable, but also meaningful for training prescription, rehabilitation, performance monitoring, injury prevention, or clinical decision-making. Similarly, Holmberg [12] emphasises that recurring debates on statistical significance reflect deeper conceptual misunderstandings about probability and inference, and calls for clearer, actionable guidance to support sound decision-making in applied research settings. In parallel, recent methodological work on interpretable machine learning highlights the need for the growing uptake of data-driven models to be accompanied by transparent validation, interpretability, and clear communication of model uncertainty and limitations [13].

The aim of this methodological narrative review is to provide a methodological synthesis of common statistical challenges in exercise and sport sciences and to propose a decision-oriented framework for model selection and reporting. The intended audience includes

applied exercise and sport science researchers, early-career investigators, doctoral students, journal reviewers, and interdisciplinary collaborators who have basic statistical knowledge but require structured guidance for aligning research questions, data structures, and analytical strategies. By framing statistical modelling as a sequence of explicit analytical decisions, this contribution seeks to support researchers in selecting appropriate methods that are aligned with both substantive research questions and empirical data characteristics. Rather than introducing new statistical methods or presenting a newly validated statistical model, this contribution focuses on improving applied statistical reasoning by translating established methodological principles into a practical decision-oriented framework for exercise and sport science research.

Methodological approach

This paper was designed as a methodological narrative review with applied sport-science examples and a practical focus on supporting decision-making. Advanced approaches, such as penalised regression, machine learning, and dimensionality reduction, are discussed at a conceptual level, with the aim of clarifying when they may be appropriate and when expert statistical collaboration may be required. The purpose was not to quantify the prevalence of statistical errors or to provide an exhaustive mapping of all available literature, but to synthesise key methodological issues that repeatedly arise in exercise and sport sciences and to translate them into practical decision-oriented guidance for researchers, reviewers, and practitioners.

A targeted literature search was conducted to identify methodological contributions relevant to statistical modelling, inference, prediction, validation, reproducibility, and applied data analysis in exercise, sport, health, and biomedical research. Searches were performed in major scientific databases and academic search engines, including PubMed, Scopus, Web of Science, and Google Scholar. Search terms included combinations of keywords such as “exercise science”, “sport science”, “statistical modelling”, “mixed-effects models”, “multilevel models”, “machine learning”, “prediction models”, “validation”, “calibration”, “overfitting”, “estimand”, “reproducibility”, “robust statistics”, and “statistical inference”.

The selection of sources followed a purposive and concept-driven approach. Priority was given to: (i) methodological papers and textbooks that define core statistical principles; (ii) sport- and exercise-science papers highlighting applied analytical problems; (iii) recent methodological contributions published between 2020

and 2025; and (iv) papers relevant to the construction of a practical decision framework for model selection and reporting. Particular attention was given to recent exercise- and sport-science-specific methodological literature addressing transparency, reproducibility, statistical power, statistical collaboration, questionable research practices, reporting bias, and the need for improved analytical planning [1–6]. Foundational references published before 2020 were retained when they represented widely accepted methodological standards or concepts that remain central to contemporary statistical practice. Sources were included when they contributed directly to one or more of the review aims: defining methodological problems, clarifying statistical concepts, illustrating applied sport-science challenges, or supporting the development of the decision-oriented framework. Sources were not prioritised when they were unrelated to statistical modelling, did not address methodological decision-making, or provided only narrow technical details without relevance to applied exercise and sport science research. Although no standardised extraction form was employed, the selected sources were consistently mapped onto the main conceptual domains of the review: study design and data structure, model selection and overfitting, assumptions and interpretability, validation and reporting, and practical application of the decision-oriented framework. To clarify the screening procedure, potentially relevant sources were first assessed for methodological relevance to statistical modelling and applied exercise, sport, health, or biomedical research. Eligible full texts were then retained when they contributed conceptually to at least one review domain or to the construction of the decision-oriented framework.

The synthesis was therefore critical and integrative rather than exhaustive. It did not aim to restate statistical principles in isolation, but to identify recurrent mismatches between research objectives, data structures, model assumptions, validation practices, and reporting standards in exercise and sport sciences. These recurring mismatches were used to define the main decision points of the proposed framework. The synthesis was organised around four conceptual domains: study design and data structure, model selection and overfitting, assumptions and interpretability, and model validation and reporting. These domains were selected because they correspond to recurrent decision points in applied exercise and sport science research. The decision-oriented framework and accompanying tables were then developed by mapping these domains onto common analytical scenarios, including explanatory and predictive aims, continuous and categorical out-

comes, repeated or clustered data, and different levels of model complexity.

Because this manuscript was designed as a methodological narrative review rather than as a systematic review or meta-analysis, no formal study-level risk-of-bias tool was applied. However, the selected literature was assessed qualitatively for methodological relevance, conceptual rigour, field-specific applicability, consistency with established statistical principles, and contribution to the development of the decision-oriented framework. The review therefore uses a transparent narrative synthesis to integrate methodological literature, reporting recommendations, and applied examples into practical guidance for exercise and sport sciences. Because this was not a systematic review, the conclusions should be interpreted as a methodological synthesis based on the triangulation of established statistical principles, reporting guidance, and recurrent applied problems, rather than as evidence weighted by formal risk-of-bias scores or quantitative synthesis.

Common methodological challenges in exercise and sport sciences

Study design and data structure

Research in exercise and sport sciences encompasses a wide range of study designs, including randomised controlled trials, quasi-experimental interventions, observational cohort studies, and cross-sectional analyses. Each design implies specific assumptions about causal inference, dependency structures, and potential sources of bias. These design-related implications are not always explicitly acknowledged during the statistical analysis phase, which may lead to the use of models only partially compatible with the underlying data-generating process. In many cases, statistical models are selected without explicit consideration of these design features, increasing the risk of biased inference and suboptimal analytical choices.

Small sample sizes are particularly common in intervention studies involving athletes or clinical populations, leading to limited statistical power and increased uncertainty in parameter estimates [1]. In such contexts, conventional asymptotic approximations may be unreliable, and the risk of both type I and type II errors increases substantially. Additionally, many datasets involve repeated measures over time, nested structures (e.g., athletes within teams), or clustered sampling designs. In longitudinal or time-dependent designs, repeated measurements may also be affected by temporal artefacts, circadian or biological rhythms, and

time-of-day effects, which should be distinguished from genuine intervention or performance-related changes through appropriate temporal control and modelling strategies [14]. Concrete measurement-derived examples further illustrate this issue. Force-plate assessments, for instance, often generate repeated trials within the same athlete and may be influenced by sampling frequency, filtering procedures, movement execution, device calibration, and signal-processing choices. In push-up power assessment, methodological issues related to force-plate use can affect the interpretation of force-derived outcomes and should therefore be considered prior to statistical modelling [15]. These complexities may increase when kinetic signals are integrated with mathematical movement models, because derived variables such as flight time, displacement, height, and mechanical power depend not only on raw force data, but also on the assumptions and parameterisation of the biomechanical model used to quantify movement [16]. From an analytical perspective, repeated force-plate trials should not be treated as fully independent observations. When these dependencies are ignored and standard regression or analysis of variance models are applied, standard errors and confidence intervals can be biased; hence, mixed-effects or related approaches are required to account for clustering and repeated measures [17].

Violations of fundamental assumptions, such as the independence of observations, remain among the most common and consequential methodological errors in applied sport science research. This issue is particularly salient in longitudinal training studies, school-based physical activity interventions, and team sport analyses, where hierarchical data structures naturally induce within-unit correlation. Mixed-effects and multilevel models offer statistically appropriate solutions for handling such dependency, yet their correct application requires careful attention to underlying assumptions, which are frequently violated in practice [17].

Model selection and overfitting

Model selection in exercise and sport sciences is often driven by convention rather than explicit criteria. Multiple regression models are frequently constructed by including all available predictors or by relying on automated stepwise procedures, practices that increase the risk of overfitting and unstable estimates [18]. Overfitted models may exhibit apparently strong in-sample associations while failing to generalise to new samples or settings, thereby limiting their practical utility. These issues are particularly problematic in small samples or

when the number of predictors approaches the sample size. In predictive modelling, overfitting may also occur through data leakage when feature selection, pre-processing, dimensionality reduction, or model tuning is performed on the full dataset before cross-validation or testing. This allows information from the validation or test data to influence model development, leading to optimistically biased performance estimates. To avoid this problem, feature selection and pre-processing steps should be performed within the resampling loop or within a clearly separated training set.

A further conceptual issue concerns the frequent lack of distinction between explanatory and predictive modelling. Explanatory models aim to estimate interpretable associations or causal effects, whereas predictive models prioritise out-of-sample performance. Failure to clarify this distinction may result in inappropriate model evaluation, such as assessing predictive models solely on statistical significance rather than on predictive accuracy or calibration. Confusing these objectives can lead to inappropriate evaluation metrics and misleading conclusions [19]. A practical example is the prediction of one-repetition maximum (1-RM) from submaximal loads or body-mass-related indicators in strength and conditioning research. In this case, the aim is not primarily to interpret each predictor as a causal determinant of maximal strength, but to estimate 1-RM accurately in athletes not used to develop the model. Therefore, model evaluation should focus on prediction error, calibration, and out-of-sample generalisability rather than only on the statistical significance of individual predictors. In this context, independent validation cohorts are particularly important because they test whether prediction equations retain acceptable accuracy in athletes not used for model development, thereby reducing the risk of overfitting and inflated performance estimates [20]. Recent methodological discussions emphasise that explicitly defining the estimand and clarifying the modelling objective are essential steps to avoid misinterpretation and analytical pitfalls. The International Council for Harmonisation (ICH) E9(R1) framework highlights how a precise formulation of the estimand aligns the study question, design, data collection, and analysis, thereby improving the transparency and validity of statistical inference [21]. For example, in a resistance-training intervention comparing two protocols, the estimand should specify whether the objective is to estimate the causal effect of the assigned protocol on post-intervention strength, adjusted for baseline strength, or the effect among participants who fully adhered to the programme. It should also clarify how missing follow-up

assessments, training non-adherence, or protocol deviations are handled. In this scenario, the estimand directly informs whether the model should prioritise adjusted effect estimation, per-protocol interpretation, intention-to-treat reasoning, or sensitivity analyses. Explicit articulation of the estimand is therefore a prerequisite for coherent model selection, appropriate interpretation of results, and transparent communication of study findings in exercise and sport sciences.

Assumptions, robustness, and interpretability

Statistical models rely on assumptions that are often insufficiently assessed or reported in applied research. Linearity, normality, homoscedasticity, and independence are frequently assumed rather than empirically evaluated. In exercise and sport sciences, violations of these assumptions are common due to physiological variability, ceiling or floor effects, and non-linear dose-response relationships. When violated, these assumptions may compromise both inference and prediction [22].

The growing use of machine learning models has further intensified concerns regarding interpretability. While complex algorithms may improve predictive accuracy, their practical value in clinical and sport settings depends on the ability to explain and contextualise results. Without adequate interpretability, statistical models may be difficult to translate into actionable recommendations for practitioners, coaches, and clinicians. Recent methodological work emphasises the importance of model interpretability, transparency, and robustness as central components of responsible data analysis [13].

Statistical modelling approaches in exercise and sport sciences

Classical statistical models

Classical statistical models remain foundational in exercise and sport sciences. Linear and generalised linear models allow the estimation of associations between predictors and outcomes under clearly defined assumptions. Mixed-effects and multilevel models extend these approaches to accommodate hierarchical data structures, repeated measures, and individual variability [23]. These models offer a favourable balance between interpretability and flexibility, making them particularly suitable for explanatory research questions.

Mediation and moderation analyses are frequently used to investigate mechanisms and effect modification in intervention studies. However, causal interpretation of mediation effects requires stronger assumptions than ordinary association models. These include correct model specification, adequate sample size, appropriate temporal ordering of exposure, mediator, and outcome, no unmeasured exposure-outcome confounding, no unmeasured exposure-mediator confounding, no unmeasured mediator-outcome confounding, and no mediator-outcome confounders affected by the exposure [24]. These assumptions are often difficult to satisfy in observational sport and exercise studies, where training exposure, motivation, baseline fitness, fatigue, recovery status, and contextual factors may jointly influence both mediators and outcomes. Measurement error in the mediator is another important concern, because imprecise or unreliable mediator assessment may bias both direct and indirect effect estimates and weaken the interpretation of mediation pathways [24]. When these assumptions and measurement-related requirements are not met, mediation results should be interpreted as exploratory or associational rather than as evidence of a causal mechanism.

Advanced and data-driven approaches

Penalised regression techniques, such as ridge regression and the least absolute shrinkage and selection operator (LASSO), offer solutions to multicollinearity and high-dimensional predictor spaces. LASSO is a regularisation method that simultaneously performs variable selection and coefficient shrinkage by imposing an L1 penalty on regression coefficients, thereby reducing model complexity and improving generalisability. These approaches, together with related methods such as elastic net regression, are extensively described in modern statistical learning resources [25] and are particularly useful in contexts involving numerous correlated performance or physiological indicators. For example, in multi-sensor wearable technology datasets, researchers may collect numerous correlated predictors, including global positioning system (GPS)-derived distance, speed zones, accelerations, decelerations, heart-rate responses, training load indicators, and ratings of perceived exertion. In such settings, LASSO or elastic net regularisation can reduce high-dimensional collinearity and overfitting by shrinking unstable coefficients and selecting a more parsimonious subset of predictors for prediction or exploratory modelling.

In predictive modelling, penalisation alone does not prevent all sources of optimistic performance estima-

tion. A particularly important risk is data leakage, which occurs when feature selection, pre-processing, dimensionality reduction, or model tuning is performed using information from the full dataset before cross-validation or testing. In such cases, the validation procedure is no longer independent, and predictive performance may be substantially overestimated [26]. Therefore, feature selection and pre-processing steps should be conducted within the resampling loop or within a clearly separated training set.

Similarly, high-dimensional biological data in sport science may arise from electromyography, kinematic time-series, and musculoskeletal modelling outputs. These datasets often contain correlated signals across muscles, joints, movement phases, or repeated trials, making dimensionality reduction and advanced modelling essential for extracting interpretable neuromuscular patterns. For example, Tajik *et al.* [27] used non-negative matrix factorisation to analyse shoulder muscle synergies during the badminton forehand overhead smash, integrating electromyography-derived signals with musculoskeletal modelling outputs. This type of application illustrates how dimensionality reduction can transform complex multidimensional sport data into interpretable structures relevant to performance analysis, injury prevention, and rehabilitation.

Supervised machine learning models have gained popularity for prediction and classification tasks in sport performance and health risk assessment. Nevertheless, their application requires careful validation and an explicit trade-off between predictive performance and interpretability [7]. In applied sport science, prioritising predictive accuracy without adequate interpretability may limit the explanatory and translational value of findings. This issue is particularly relevant for athlete sensor technologies and artificial intelligence applications, where algorithmic transparency, standardised assessment procedures, and clear reporting of model inputs, pre-processing steps, and decision rules are essential for interpreting and comparing results across studies [28]. Recent guidance further emphasises that predictive performance should not be evaluated solely in terms of discrimination, but also in terms of calibration and clinical or practical utility, particularly when models are intended to support decision-making in health and sport settings [26].

These advanced approaches should not be treated as interchangeable technical options. Penalised regression is mainly useful for shrinkage, multicollinearity, and high-dimensional predictors; machine learning is primarily suited to prediction and classification when validation, calibration, leakage prevention, and inter-

pretability are addressed; Bayesian estimation may be useful when probabilistic interpretation, prior information, hierarchical structures, or shrinkage are relevant; and causal modelling requires explicit assumptions about the estimand, confounding structure, temporal ordering, and missing data.

Model validation and performance assessment

Robust statistical practice requires both internal and external validation. Cross-validation and bootstrap procedures provide estimates of model stability and predictive performance, while sensitivity analyses assess robustness to analytical choices. External validation, although less frequently performed, is essential for evaluating generalisability across populations and contexts. Importantly, reliance on *p*-values alone is insufficient; effect sizes, confidence intervals, and calibration metrics should be routinely reported [29].

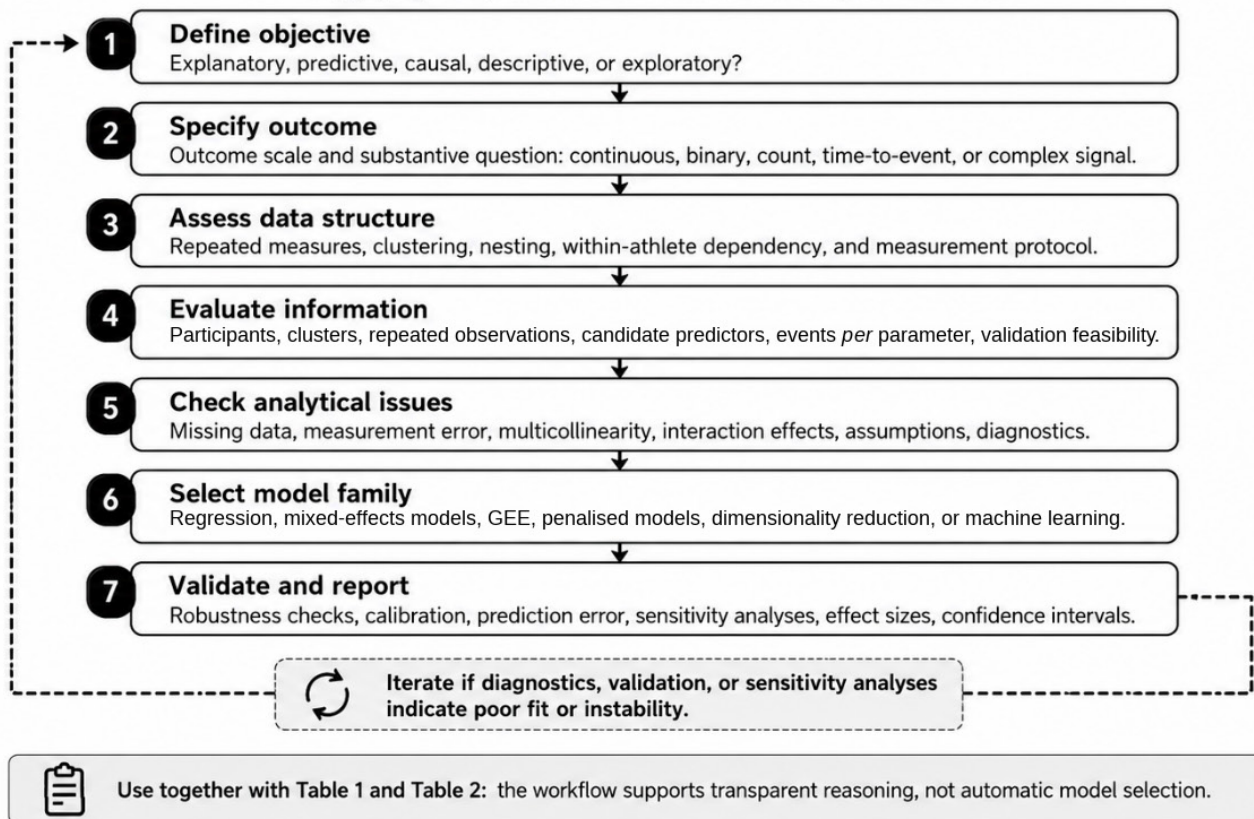
A decision-oriented framework for statistical modelling

Statistical modelling in exercise and sport sciences is often treated as a purely technical step performed after data collection. However, inappropriate or implicit analytical choices frequently represent a major source of bias, limited interpretability, and reduced reproducibility. To address this issue, the present framework conceptualises statistical modelling as a decision-oriented process, in which analytical choices are explicitly derived from substantive research objectives and empirical data characteristics. The framework should be understood as a methodological decision aid rather than as a newly validated statistical model. Its contribution lies in organising established statistical principles into an explicit sequence of analytical decisions tailored to common problems in exercise and sport sciences.

The proposed framework operationalises statistical modelling as a sequence of explicit analytical decisions. Rather than starting from available software tools or conventional tests, researchers are encouraged to first clarify the primary aim of the analysis (explanatory *vs.* predictive), followed by the definition of the outcome scale, the identification of dependency or clustering in the data, and an evaluation of information content relative to model complexity. These steps collectively define the estimand and constrain the set of statistically coherent modelling options, reducing analytical flexibility and post-hoc decision-making.

Decision-Oriented Statistical Workflow

Applying the framework in exercise and sport sciences



The workflow illustrates the main analytical steps involved in model selection and reporting: (1) define the study objective; (2) identify the outcome type and substantive question; (3) assess the data structure, including repeated measures, clustering, and dependency; (4) evaluate information content and model complexity; (5) consider key analytical issues such as missing data, measurement error, multicollinearity, assumptions, and interaction effects; (6) select an appropriate model family; and (7) perform validation, diagnostics, and transparent reporting. The workflow is intended as a practical decision aid and should be used together with Table 1 and Table 2. GEE – generalised estimating equations.

Figure 1. Step-by-step workflow for applying the decision-oriented statistical framework in exercise and sport sciences

To make the framework more operational, Figure 1 presents a step-by-step workflow for applying decision-oriented statistical reasoning in exercise and sport sciences. The workflow is intended to guide researchers from the initial definition of the analytical objective to model selection, validation, and reporting. Rather than replacing statistical judgment, the figure illustrates the main sequence of decisions that should be made explicitly when choosing and justifying a statistical approach. Table 1 and Table 2 should be interpreted as complementary tools that provide more detailed guidance for specific analytical scenarios and methodological options.

In practical terms, the workflow encourages researchers to move from the research question to the statistical model through a transparent sequence of justified decisions, rather than selecting analytical methods only on the basis of convention or software availability.

Table 1 presents a decision-oriented framework for statistical model selection in exercise and sport sciences. The framework does not classify sample size using rigid labels such as “small”, “moderate”, or “large.” Instead, it emphasises information content, including the number of participants, repeated observations, clusters, candidate predictors, events *per* parameter, and the feasibility of validation. The table also incorporates additional analytical considerations that are essential in real research settings, including missing data, measurement error, multicollinearity, interaction effects, model assumptions, diagnostics, and robustness checks. The framework is not intended to replace specialised statistical expertise, but to make explicit the main analytical decisions that are often left implicit in applied exercise and sport science manuscripts. Its added value is therefore primarily editorial and methodological: it provides a structured basis for planning, justifying, reviewing, and reporting statistical analyses.

Table 1. Decision-oriented framework for model selection in exercise and sport sciences

Study objective	Outcome / data structure	Key analytical issues	Recommended model family	Sample size/information guidance	Validation and diagnostics
Explanation/ effect estimation	continuous outcome, independent observations	linearity, residual distribution, influential observations, covariate adjustment	linear regression/ ANCOVA	power analysis based on expected effect size; avoid excessive predictors relative to sample size	residual diagnostics, influence analysis, effect sizes and confidence intervals
Explanation/ effect estimation	continuous outcome, repeated or clustered data	within-subject correlation, clustering, unequal follow-up, missing repeated observations	linear mixed-effects models	consider number of participants, number of repeated measures, and number of clusters; avoid overfitting random effects	random-effects structure, residual diagnostics, sensitivity analyses
Explanation/ effect estimation	binary or count outcome	non-normal outcome distribution, overdispersion, sparse data	logistic, Poisson, negative binomial, or generalised linear mixed models	consider events per parameter, including conventional events-per-variable heuristics for logistic models; assess overdispersion and avoid unstable models with sparse events	calibration, goodness-of-fit, overdispersion checks, confidence intervals
Causal/ mechanism-oriented analysis	exposure, mediator, outcome, observational or intervention data	temporal ordering, confounding, mediator-outcome confounding, interaction effects, missing data, assumption reporting	mediation/ moderation models; causal modelling when assumptions are explicit	requires sufficient sample size for indirect effects and clear temporal structure	sensitivity analysis for confounding, effect decomposition, transparent assumption reporting
Prediction	continuous or categorical outcome with multiple predictors	multicollinearity, high-dimensional predictors, overfitting, model instability	penalised regression, elastic net, supervised machine learning	consider predictor-to-observation ratio, events-per-variable for categorical outcomes, model optimism, and feasibility of internal or external validation	cross-validation, bootstrap validation, calibration, discrimination, prediction error
Prediction/ classification	complex sensor, wearable, GPS, EMG, or biomechanical data	high dimensionality, signal processing, correlated predictors, interpretability	dimensionality reduction, penalised models, machine learning, interpretable models	requires sufficient observations relative to feature space; feature reduction or regularisation is often necessary	internal validation, external validation when possible, feature importance, calibration, interpretability checks
Population-average inference	repeated or clustered data where marginal effects are of interest	correlated responses, cluster-level inference, population-average interpretation	generalised estimating equations	requires adequate number of clusters for stable estimation	robust standard errors, working correlation structure, sensitivity checks

ANCOVA – analysis of covariance, EMG – electromyography, GPS – global positioning system

Sample-size adequacy should not be interpreted through universal labels such as “small”, “moderate”, or “large”. Instead, model complexity should be justified through *a priori* power analysis for explanatory models, effective sample size, number of clusters, number of repeated observations, number of candidate predictors, and events *per* parameter for binary or sparse event-based outcomes, including conventional events-*per*-variable heuristics and, where possible, context-specific prediction-model sample-size calculations. For predictive models, feasibility of internal or external validation should also be considered.

For binary outcomes and categorical prediction models, events-*per*-variable criteria may provide useful preliminary heuristics, such as the conventional 10 events-*per*-variable rule for logistic regression, although more recent guidance recommends context-specific sample-size calculations that account for outcome prevalence, anticipated model performance, shrinkage, and validation needs [30, 31].

The framework emphasises that statistical model selection cannot be based solely on outcome type or conventional sample-size labels. Instead, researchers should evaluate whether the available data contain enough information to support the planned model complexity. For example, a study with a modest number of participants but several repeated measurements may support a mixed-effects model, whereas a prediction

Table 2. Common analytical scenarios and contemporary statistical options in exercise and sport sciences

Objective	Data characteristics	Classical approach	Non-parametric option	Robust/contemporary alternatives
Independent group comparison	continuous outcome, two or more groups	<i>t</i> -test / ANOVA	Mann–Whitney / Kruskal–Wallis	Welch <i>t</i> -test / Welch ANOVA, bootstrap confidence intervals, trimmed means
Paired comparison	continuous pre-post or paired data	paired <i>t</i> -test	Wilcoxon signed-rank test	bootstrap confidence intervals for mean or median change; robust paired estimators
Repeated measures	continuous repeated observations	repeated-measures ANOVA	Friedman test	linear mixed-effects models, robust mixed models, Bayesian hierarchical models
Association	continuous variables	Pearson correlation / linear regression	Spearman correlation	robust correlation, robust regression, bootstrap confidence intervals
Binary outcome	event / non-event outcome	logistic regression / χ^2 test	Fisher's exact test	penalised logistic regression, Bayesian logistic regression, calibration assessment
Count outcome	counts, rates, event frequency	Poisson regression	–	negative binomial regression, zero-inflated models, generalised linear mixed models
High-dimensional sport data	wearables, GPS, EMG, force-plate or biomechanical features	multiple regression	–	penalised regression, dimensionality reduction, machine learning, Bayesian shrinkage models
Prediction / classification	continuous or categorical outcome	regression / classification models	–	cross-validation, bootstrap validation, calibration, decision-curve or utility-based assessment

ANOVA – analysis of variance, EMG – electromyography, GPS – global positioning system

Parametric procedures assume adequate model assumptions, including approximate residual normality for linear models, homogeneity of variance where appropriate, and independence or correctly modelled dependency according to the study design. Non-parametric tests are useful under severe non-normality, ordinal data, small samples, or strong outliers, but they are not the only solution to assumption violations. Robust statistics, bootstrapping, generalised models, mixed-effects models, and Bayesian estimation may provide more informative alternatives depending on the research question, data structure, and target estimand. Where feasible, hypothesis tests should be accompanied by effect sizes and confidence intervals, and robustness should be assessed through graphical diagnostics, sensitivity analyses, and transparent reporting of model assumptions.

model with many candidate predictors may require penalisation or dimensionality reduction even when the total sample size appears acceptable. Similarly, the presence of missing data, measurement error, multicollinearity, or sparse outcomes may require alternative model specifications or sensitivity analyses. Thus, the framework is intended to support transparent reasoning rather than replace statistical judgment.

By formalising statistical modelling as a sequence of justified decisions, this framework supports transparent reporting, facilitates peer review, and enhances reproducibility. Moreover, it provides a common methodological language for interdisciplinary research

teams, improving communication between statisticians, exercise scientists, clinicians, coaches, and practitioners.

To complement the model-selection framework, Table 2 summarises classical, non-parametric, robust, and model-based alternatives for common analytical scenarios in exercise and sport sciences. The purpose is to avoid presenting non-parametric tests as the only alternative to violated parametric assumptions. In contemporary applied statistics, assumption violations can often be addressed using robust methods and resampling procedures, such as trimmed means, robust estimators, and bootstrap confidence intervals [32], as well as model-based approaches that explicitly account

for outcome distribution, dependency structure, and uncertainty [33]. Bayesian estimation may also be considered when probabilistic interpretation, prior information, or shrinkage-based reasoning are relevant to the research question [34].

Table 2 should be interpreted as a practical decision aid, not as a rigid algorithm for test selection. In particular, the choice between parametric and non-parametric methods should not be based exclusively on formal normality tests, which may be underpowered in small samples and overly sensitive in moderate or large samples. Researchers should combine graphical diagnostics, distributional summaries, sensitivity analyses, and estimation-focused reporting, including effect sizes and confidence intervals where appropriate. When assumptions are violated, possible solutions include non-parametric procedures, robust estimators, bootstrap confidence intervals, generalised linear models, mixed-effects models, or Bayesian estimation, depending on the analytical objective and data structure [32–34]. This broader perspective is especially relevant in exercise and sport sciences, where repeated measurements, clustered observations, outliers, skewed physiological variables, and high-dimensional sensor-derived data are common.

Applied exercise and sport science examples: operationalising the framework

To address the practical application of the proposed decision-oriented framework, this section presents four applied examples from exercise and sport sciences, ranging from sport performance and biomechanics to clinical exercise interventions. These examples are not intended as new empirical analyses, but as applied scenarios showing how the framework can be used to connect research objectives, data structure, measurement characteristics, and model selection.

Force-plate-derived biomechanical data

Force plates are widely used in sport science to assess power, force production, jump performance, and movement mechanics. However, force-plate data often involve repeated trials, high-frequency measurements, and technical sources of variability related to sampling frequency, filtering, signal processing, movement execution, and device calibration. Consequently, observations derived from repeated trials within the same athlete cannot be treated as fully independent.

For example, when assessing push-up power through force plates, methodological issues related to measure-

ment setup and data processing may directly affect the reliability and interpretation of the derived outcomes [15]. When force-plate signals are combined with analytical or mathematical movement models, additional modelling assumptions may influence the estimation of flight time, arc displacement, maximum height, and mechanical power output, as illustrated by rigid-body pendulum modelling of plyometric push-up biomechanics [16]. Within the proposed framework, the researcher should first clarify whether the objective is explanatory, such as estimating the effect of a training intervention on force output, or predictive, such as classifying athletes according to performance profiles. If repeated trials are collected from each participant, mixed-effects models may be preferable to standard linear regression or repeated univariate tests because they can account for within-athlete correlation and inter-individual variability.

This example illustrates that the data structure is not determined only by the statistical question, but also by the measurement technology and protocol. Instrument-derived variables should therefore be modelled in a way that reflects the dependency structure generated by the assessment procedure.

Predicting maximal strength from submaximal performance

Strength and conditioning research often aims to predict one-repetition maximum (1-RM) from submaximal loads, body mass, or repetition-based performance indicators. In this context, the primary aim is typically predictive rather than explanatory: the model is expected to estimate future or unobserved maximal strength values with acceptable accuracy in athletes not included in the model development sample.

Recent applied research has examined the prediction of 1-RM and 4–6 repetition maximum performance in free-weight bench press and Smith machine squat based on body mass in male athletes [20]. Within the proposed framework, this type of study requires explicit separation between model development and model evaluation. Model performance should not be judged only by statistical significance of predictors, but by predictive accuracy, calibration, and out-of-sample performance. Internal validation through resampling procedures is useful for estimating optimism, but independent validation cohorts remain essential when prediction equations are intended for practical use in new athletes or settings. External validation helps verify whether the model maintains acceptable prediction error, calibration, and generalisability beyond the development sample.

This example clarifies the practical difference between explanatory and predictive modelling. A regression coefficient may be statistically significant and theoretically interpretable, but the resulting model may still perform poorly when used to predict strength in new athletes. Therefore, predictive sport-science models should be evaluated with metrics that reflect prediction error, calibration, and practical usefulness.

Multi-sensor electromyography and musculoskeletal modelling

Another example concerns high-dimensional biomechanical and neuromuscular data, such as electromyography, kinematic time-series, and musculoskeletal modelling outputs. These data often include many correlated variables measured across muscles, joints, movement phases, or trials. In such contexts, classical univariate analyses may be inadequate because they ignore multicollinearity, temporal dependency, and the multidimensional structure of the signal.

Recent work integrating electromyography and musculoskeletal modelling during the badminton forehand overhead smash illustrates this complexity. Tajik *et al.* [27] used non-negative matrix factorisation to quantify shoulder muscle synergies and compared experimentally derived electromyographic synergies with musculoskeletal model outputs. Such an approach exemplifies how dimensionality reduction methods can transform complex multi-muscle activation patterns into interpretable lower-dimensional structures.

Within the proposed framework, the modelling strategy should be guided by the objective of the analysis. If the aim is to explain coordination patterns, dimensionality reduction methods, such as non-negative matrix factorisation, may be appropriate. If the aim is to predict performance or injury risk from multi-sensor data, penalised regression or supervised machine learning may be considered, provided that validation, calibration, and interpretability are explicitly addressed. This example shows that advanced modelling is not simply a matter of using complex algorithms, but of matching the analytical strategy to the biological structure of the data and the practical question being asked.

Clinical exercise interventions in non-communicable diseases

The framework can also be applied to clinical exercise science, particularly to exercise interventions in populations living with non-communicable diseases. For example, in a trial evaluating a multicomponent

exercise programme for adults with cardiovascular disease, type 2 diabetes, obesity, or cancer-related functional limitations, the primary objective may be explanatory, predictive, or both. If the aim is to estimate the intervention effect on functional capacity, fatigue, glycaemic control, blood pressure, or quality of life, the model should account for baseline values, repeated follow-up assessments, missing data, adherence, medication changes, and relevant clinical covariates. These choices should be aligned with the target estimand and the intended interpretation of the intervention effect [21]. Linear mixed-effects models or generalised mixed models may be appropriate when outcomes are repeatedly measured or when participants are clustered within centres, exercise groups, or rehabilitation settings [17].

In this context, the decision-oriented framework helps clarify whether the analysis is intended to estimate an average intervention effect, identify moderators of response, predict individual adaptation, or support clinical decision-making. It also emphasises the need to distinguish intention-to-treat and per-protocol questions, handle missing follow-up data transparently, and report effect sizes, confidence intervals, clinically meaningful changes, and sensitivity analyses. When prediction is involved, calibration, prediction error, internal validation, and, where possible, external validation should also be considered [26, 29]. This example extends the framework beyond sport performance and biomechanics to clinical exercise interventions in which statistical decisions may influence rehabilitation, prevention, and the long-term management of non-communicable diseases.

Discussion and future perspectives

The methodological issues highlighted in this review article reflect a broader tension in exercise and sport sciences: the rapid growth of available analytical tools has not always been accompanied by parallel improvements in methodological reasoning, transparency, and reproducibility. The increasing use of machine learning and high-dimensional datasets affords opportunities for more accurate prediction and richer modelling of complex phenomena, but also increases the risk of overfitting, misinterpretation, and limited generalisability when validation and reporting practices are inadequate. In applied domains, where research findings may influence training prescription, rehabilitation protocols, and public health recommendations, these risks carry direct practical implications. In this context, the proposed framework and decision

tables offer pragmatic tools to translate methodological principles into day-to-day analytical practice.

A key contribution of this paper is the proposal of a decision-oriented framework that translates methodological principles into a sequence of explicit analytical decisions. By aligning model choice with study objective, outcome type, data structure, and information content, the framework is intended to support coherent statistical reasoning and facilitate communication within interdisciplinary teams. This is particularly relevant in exercise and sport sciences, where collaboration between methodologists, clinicians, coaches, and sport scientists requires a shared language to interpret quantitative evidence. Embedding statisticians or methodological experts early in the research process, rather than only at the analysis stage, can improve study design efficiency, reduce analytical flexibility, and strengthen the credibility of results. This recommendation is consistent with the call by Sainani *et al.* [4] to increase statistical collaboration in sports science, sport and exercise medicine, and sports physiotherapy. The subsequent responses to this call further reinforce that improving statistical practice should not be viewed only as an individual researcher responsibility, but as a shared responsibility involving authors, statistical collaborators, reviewers, editors, and journals.

From a future-oriented perspective, three priorities emerge. First, methodological education in exercise and sport sciences should increasingly emphasise “decision literacy” rather than software-driven analysis: researchers should be trained to justify model selection, interpret uncertainty, and recognise common pitfalls such as model fishing and post-hoc hypothesis construction. Second, greater adoption of structured analytical planning (*e.g.*, brief statistical analysis plans, pre-registration where feasible, and transparent reporting of deviations) could reduce waste and improve reproducibility. Third, modern applied research should incorporate stronger validation culture, including internal validation as a minimum standard for predictive models and external validation whenever generalisability claims are made.

To make decision literacy more operational for journals, editorial boards, and peer reviewers, Table 3 provides a practical checklist that can be used during manuscript evaluation. This checklist is not intended to replace expert statistical review, but to support a more transparent assessment of whether the main analytical decisions are justified, reproducible, and aligned with the study objective.

Table 3 provides a practical mechanism for translating decision literacy into routine editorial practice.

It encourages reviewers to evaluate not only whether a statistical technique is technically acceptable, but also whether it is appropriate for the research objective, data structure, and intended interpretation of the findings.

The relevance of decision-oriented statistical modelling is particularly evident in biomechanics. Biomechanical research frequently involves force-plate signals, kinematic time-series, electromyography, wearable sensors, and musculoskeletal modelling outputs, which generate complex, high-dimensional, and dependency-rich data structures. In this sub-discipline, robust statistical reasoning is not only a methodological requirement, but also a prerequisite for translating data into valid conclusions about performance optimisation, injury-risk mitigation, and rehabilitation planning. Recent perspectives on biomechanics emphasise that advances in hardware, software, sensor technologies, and modelling approaches can improve sports performance, reduce injury risk, and support athlete rehabilitation, but only when analytical methods are sufficiently rigorous, transparent, and interpretable [35]. Therefore, methodological education in exercise and sport sciences should include sub-discipline-specific training that helps researchers connect statistical decisions with the biological, biomechanical, and practical meaning of the data.

It should also be acknowledged that this methodological narrative review does not provide a quantitative estimate of the prevalence or severity of statistical problems in exercise and sport sciences. The issues discussed in this paper should therefore be interpreted as recurrent methodological concerns identified through the selected literature and applied examples, rather than as findings from a systematic meta-research analysis. Future studies could extend this work by conducting systematic methodological audits or meta-research investigations to quantify the frequency of small-sample designs, assumption violations, overfitting, inadequate validation, and incomplete reporting across exercise and sport science journals.

The proposed framework should therefore be interpreted as a practical methodological synthesis and decision aid, and not as an empirically validated prediction tool or a new statistical model. Accordingly, claims regarding improved reproducibility and analytical reasoning should be interpreted as the intended methodological purpose of the framework, rather than as empirically demonstrated effects. Future studies could extend this work by testing the framework through simulation studies, real-data applications, or evaluations of its impact on statistical reporting quality in

Table 3. Editorial and peer-review checklist for statistical decision literacy in exercise and sport sciences

Checklist item	Purpose for editorial and peer-review assessment
Has the primary analytical objective been clearly defined as explanatory, predictive, causal, descriptive, or exploratory?	Clarifies whether the analysis is intended to estimate effects, predict outcomes, describe patterns, or explore hypotheses.
Is the selected statistical model aligned with the study design, outcome type, data structure, and target estimand?	Ensures coherence between the research question and the analytical strategy.
Are repeated measurements, clustering, nesting, or within-athlete dependencies explicitly modelled or justified?	Prevents biased standard errors and misleading inference due to ignored dependency structures.
Is sample-size adequacy justified through power analysis, effective sample size, number of clusters, repeated observations, events <i>per</i> parameter, or predictor-to-observation considerations?	Assesses whether the available data support the planned model complexity.
Are missing data, protocol deviations, measurement error, and influential observations described and handled transparently?	Improves interpretability and reduces hidden analytical bias.
Are model assumptions assessed through diagnostics, graphical inspection, robustness checks, or sensitivity analyses?	Evaluates whether the reported findings are robust to plausible analytical choices.
Are effect sizes, confidence intervals, uncertainty estimates, and practical relevance reported alongside <i>p</i> -values?	Promotes estimation-focused interpretation rather than dichotomous significance-based conclusions.
For predictive models, are internal validation, calibration, prediction error, and, where possible, external validation reported?	Assesses whether predictive performance is likely to generalise beyond the development sample.
Are complex or machine-learning models accompanied by adequate interpretation, validation, and explanation of practical utility?	Reduces the risk of black-box modelling without translational value for sport science practice.
Are analytical decisions, deviations from the planned analysis, and limitations clearly reported?	Allows readers, reviewers, and editors to evaluate credibility, transparency, and reproducibility.
Are the statistical software, software version, packages or libraries, package versions when relevant, random seed numbers, and main pre-processing or analysis scripts reported or made available when feasible?	Supports computational reproducibility and allows readers, reviewers, and editors to understand the analytical environment and workflow used to generate the results.
Where relevant, are missing-data mechanisms, multiplicity, model instability in small samples, feature-engineering steps, fairness or bias in predictive modelling, and alignment with reporting guidelines explicitly addressed?	Encourages reviewers to evaluate methodological issues that are increasingly relevant in sport-science analytics but may not apply uniformly to all study designs.

exercise and sport science manuscripts. Future methodological work should also provide reproducible code, real-data demonstrations, diagnostic interpretation, and comparisons among alternative modelling strategies. The examples included in the present manuscript should therefore be interpreted as decision-mapping scenarios rather than as empirical demonstrations with model outputs, diagnostic plots, or comparative performance analyses.

Finally, the field would benefit from a clearer distinction between explanatory and predictive goals in study reporting. Explanatory models are central for

mechanism-oriented science and causal interpretation, whereas predictive models support risk stratification and decision support. This distinction should also be operationalised through established reporting guidelines, including Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) for observational studies and Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) for prediction model studies, because these frameworks help researchers report study design, model development, validation, and interpretation more transparently [36, 37]. Both

approaches are valuable, but they require different evaluation criteria and reporting standards. Recognising this distinction explicitly is critical to prevent the common error of interpreting predictive accuracy as evidence of causality or interpreting statistically significant coefficients as predictive tools.

Conclusions

Improving statistical practice in exercise and sport sciences requires a shift from technique-driven analysis to decision-oriented modelling and transparent reporting. The methodological challenges discussed in this paper, small sample sizes, dependency structures, analytical flexibility, and interpretability, are not merely technical issues, but determinants of the credibility and transferability of scientific evidence. The proposed framework and decision tables are intended to provide pragmatic tools for aligning model selection with research objectives and data characteristics, reducing preventable analytical errors, and supporting interdisciplinary communication.

Future progress will depend on the systematic integration of methodological expertise into research teams, the routine use of validation and robustness checks, and the adoption of reporting practices that emphasise estimation, uncertainty, and practical relevance. By strengthening statistical reasoning and transparency, exercise and sport sciences can produce more reliable evidence to inform clinical practice, training prescription, and public health decision-making.

Ethical approval

The conducted research is not related to either human or animal use.

Conflict of interest

The authors state no conflict of interest.

Disclosure statement

No author has any financial interest or received any financial benefit from this research.

Funding

This research received no external funding.

References

- [1] Abt G, Boreham C, Davison G, Jackson R, Nevill A, Wallace E, Williams AM. Power, precision, and sample size estimation in sport and exercise science research. *J Sports Sci.* 2020;38(17):1933–35; doi: 10.1080/02640414.2020.1776002.
- [2] Caldwell AR, Vigotsky AD, Tenan MS, Radel R, Mellor DT, Kreutzer A, Lahart IM, Mills JP, Boisgontier MP, Consortium for Transparency in Exercise Science (COTES) Collaborators. Moving sport and exercise science forward: a call for the adoption of more transparent research practices. *Sports Med.* 2020;50(3):449–59; doi: 10.1007/s40279-019-01227-1.
- [3] Lohse KR, Sainani KL, Taylor JA, Butson ML, Knight EJ, Vickers AJ. Systematic review of the use of “magnitude-based inference” in sports science and medicine. *PLOS ONE.* 2020;15(6):e0235318; doi: 10.1371/journal.pone.0235318.
- [4] Sainani KL, Borg DN, Caldwell AR, Butson ML, Tenan MS, Vickers AJ, Vigotsky AD, Warmenhoven J, Nguyen R, Lohse KR, Knight EJ, Bargary N. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. *Br J Sports Med.* 2021;55(2):118–22; doi: 10.1136/bjsports-2020-102607.
- [5] Borg DN, Barnett AG, Caldwell AR, White NM, Stewart IB. The bias for statistical significance in sport and exercise medicine. *J Sci Med Sport.* 2023; 26(3):164–68; doi: 10.1016/j.jsams.2023.03.002.
- [6] Mesquida C, Murphy J, Lakens D, Warne J. Publication bias, statistical power and reporting practices in the Journal of Sports Sciences: potential barriers to replicability. *J Sports Sci.* 2023;41(16): 1507–17; doi: 10.1080/02640414.2023.2269357.
- [7] Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods.* 2018;15(4), 233–34; doi: 10.1038/nmeth.4642.
- [8] Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol.* 2019;110:12–22; doi: 10.1016/j.jclinepi.2019.02.004.
- [9] Hernán MA, Robins JM. Causal inference: What if. Boca Raton: Chapman and Hall/CRC; 2020.
- [10] Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “ $p < 0.05$ ”. *Am Stat.* 2019;73(S1): 1–19; doi: 10.1080/00031305.2019.1583913.
- [11] McShane BB, Gal D, Gelman A, Robert C, Tackett JL. Abandon statistical significance. *Am Statist.* 2019;73(S1):235–45; doi: 10.1080/00031305.2018.1527253.
- [12] Holmberg C. Why we need to discuss statistical significance and p-values (again): addressing the underlying issue of different probability interpretations and actionable recommendations. *Nord*

- J Nurs Res. 2024; 44:1–6; doi: 10.1177/20571585241253177.
- [13] Allen GI, Gan L, Zheng L. Interpretable machine learning for discovery: statistical challenges and opportunities. *Ann Rev Stat Appl.* 2024;11:97–121; doi: 10.1146/annurev-statistics-040120-030919.
- [14] Dhahbi W, Dergaa I. Machine learning with temporal control designs for testing rhythmic specificity in chronobiology: a multivariate framework proposal for distinguishing genuine biological rhythms from temporal artifacts. *Chronobiol Int.* 2026;43(8):1118–25; doi: 10.1080/07420528.2026.2650826.
- [15] Dhahbi W, Chaouachi A, Cochrane J, Chèze L, Chamari K. Methodological issues associated with the use of force plates when assessing push-ups power. *J Strength Cond Res.* 2017;31(7):e74; doi: 10.1519/JSC.0000000000001922.
- [16] Dhahbi W. A rigid-body pendulum model for plyometric push-up biomechanics: analytical derivation and numerical quantification of flight time, arc displacement, maximum height, and mechanical power output. *Bioengineering.* 2026;13(4):445; doi: 10.3390/bioengineering13040445.
- [17] Schielzeth H, Dingemanse NJ, Nakagawa S, Westneat DF, Allogue H, Teplitsky C, Réale D, Dochtermann NA, Garamszegi LZ, Araya-Ajoy YG. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods Ecol Evol.* 2020;11(9):1141–52; doi: 10.1111/2041-210X.13434.
- [18] Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med.* 2004;66(3):411–21; doi: 10.1097/01.psy.0000127692.23278.a9.
- [19] Shmueli G. To explain or to predict? *Stat Sci.* 2010;25(3):289–310; doi: 10.1214/10-STS330.
- [20] Dhahbi W, Padulo J, Russo L, Racil G, Ltifi MA, Picerno P, Iuliano E, Migliaccio GM. 4–6 repetition maximum (RM) and 1-RM prediction in free-weight bench press and Smith machine squat based on body mass in male athletes. *J Strength Cond Res.* 2024;38(8):1366–71; doi: 10.1519/JSC.0000000000004803.
- [21] Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICHE9(R1) addendum. *BMJ.* 2024;384:e076316; doi: 10.1136/bmj-2023-076316.
- [22] Stroup WW, Ptukhina M, Garai J. Generalized linear mixed models: modern concepts, methods and applications. Boca Raton: Chapman and Hall/CRC; 2024; doi: 10.1201/9780429092060.
- [23] Asampana Asosega K, Adebajji AO, Aidoo EN, Owusu-Dabo E. Application of hierarchical/multilevel models and quality of reporting (2010–2020): a systematic review. *Sci World J.* 2024; 4658333; doi: 10.1155/2024/4658333.
- [24] Hayes AF. Introduction to mediation, moderation, and conditional process analysis: a regression-based approach. 3rd ed. Guilford Press; 2022.
- [25] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in R. 2nd ed. New York: Springer; 2021; doi: 10.1007/978-1-0716-1418-1.
- [26] Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Med Res Methodol.* 2022;22: 316; doi: 10.1186/s12874-022-01801-8.
- [27] Tajik R, Dhahbi W, Fadaei H, Mimar R. Muscle synergy analysis during badminton forehand overhead smash: integrating electromyography and musculoskeletal modeling. *Front Sports Act Living.* 2025;7:1596670; doi: 10.3389/fspor.2025.1596670.
- [28] Dhahbi W, Chamari K. The algorithmic athlete: a call to standardize assessment of sensor technologies and artificial intelligence. *Int J Sports Physiol Perform.* 2026;21(4):505–6; doi: 10.1123/ijspp.2025-0547.
- [29] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128–38; doi: 10.1097/ede.0b013e3181c30fb2.
- [30] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996;49(12): 1373–79; doi: 10.1016/S0895-4356(96)00236-3.
- [31] Riley RD, Ensor J, Snell KIE, Harrell Jr FE, Martin GP, Reitsma JB, Moons KGM, Collins G, van Smeden M. Calculating the sample size required for developing a clinical prediction model. *BMJ.* 2020;368:m441; doi: 10.1136/bmj.m441.
- [32] Wilcox RR. Introduction to robust estimation and hypothesis testing. 4th ed. Elsevier, Academic Press; 2017.
- [33] Gelman A, Hill J, Vehtari A. Regression and other stories. Cambridge University Press; 2020.

- [34] Kruschke JK. Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan. 2nd ed. Elsevier, Academic Press; 2015.
- [35] Dhahbi W. Editorial: Advancing biomechanics: enhancing sports performance, mitigating injury risks, and optimizing athlete rehabilitation. *Front Sports Act Living*. 2025;7:1556024; doi: 10.3389/fspor.2025.1556024.
- [36] von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLOS Med*. 2007;4(10):e296; doi: 10.1371/journal.pmed.0040296.
- [37] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55–63; doi: 10.7326/M14-0697.