# Multilevel Regression Modeling as a Complement to Traditional Repeated Measures Analysis of Variance in Sports Biomechanics Research

Kyle Wasserberger[1,2], William Murrah[3], Kevin Giordano[4], and Gretchen Oliver[2]

[1]Research & Development; Driveline Baseball
[2]School of Kinesiology; Auburn University
[3]Department of Educational Foundations, Leadership, & Technology; Auburn University
[4]Department of Physical Therapy; Creighton University

**Abstract**

With the increasing availability and decreasing costs of sport science technologies, longitudinal data are becoming increasingly available to the biomechanics and adjacent sport science communities. To take advantage of this emerging wealth of data and best understand how biomechanics influence health and performance over time, consideration regarding the appropriate handling of longitudinal data is warranted. Traditionally, longitudinal data in the sport sciences have most often been analyzed using repeated measures analysis of variance . However, typical analysis of variance is limited in its ability to handle missing data and unbalanced observations, two scenarios common in longitudinal designs. Alternatively, multilevel regression modeling is a flexible analysis method capable of handling these scenarios. In this paper, we hope to draw increased attention to multilevel regression modeling and lobby for its increased adaptation when examining repeated measures biomechanics data. First, we introduce the rationale and approach underlying multilevel regression models. We then present examples using simulated and real-world data to illustrate their potential application and limitations in sports biomechanics research. Lastly, we provide some recommendations, words of caution, and additional resources.

# 1 Introduction

Quantifying change over time is a central tenet for those working working with athletes of all levels. Identifying factors that influence longitudinal athletic performance is vital for evaluating the effectiveness of athlete training programs. For the biomechanist, measuring technique changes over time and elucidating the relationships between those technique changes, athlete performance, and injury risk is of particular importance. However, despite the importance placed on longitudinal athletic development in the applied sport science domain, many researchers are forced to conduct cross-sectional studies due to limited time, labor, or funding. Fortunately, technological advances continue to decrease the burdens associated with collecting longitudinal data. Given these advances, we feel increased consideration is needed regarding the appropriate means by which we analyze and draw inferences from longitudinal data.

Calls for improved statistical methodology in the sport sciences have grown in recent years [5, 10, 11, 14, 15, 19, 20, 24]. Informal skimming of recent *Sport Biomechanics* publications indicates repeated measures analysis of variance (RM·ANOVA) remains a popular choice for dealing with longitudinal biomechanics data [16, 17, 21, 28]. Beyond *Sport Biomechanics*, Google Scholar returns over 4,200 hits over the last ten years for "ANOVA" in journals associated with biomechanics compared to approximately 600 hits for "multilevel regression", "mixed-effects", or "mixed-model"[i]. Despite the greater apparent popularity of ANOVA-based methods, RM·ANOVA carries with it several important assumptions which limit its utility in applied longitudinal settings. Most notably, RM·ANOVA assumes balanced designs (equal number of repeated measures across individuals), complete cases (no missing observations within individuals), and temporal equivalence (repeated measures across individuals are equally spaced in time). While these conditions can usually be met in tightly-controlled, laboratory-based studies, applied longitudinal work often suffers from study imbalances, unforeseen disruptions to data collection schedules, and participant attrition; all of which, depending on their severity, may threaten RM·ANOVA's validity.

The assumptions surrounding RM·ANOVA present salient challenges to those working in the observational or quasi-experimental settings common in sport biomechanics research. Conversely, multilevel regression modeling is capable of handling both imbalanced designs and partially missing data, positioning it as a useful complement to more traditional repeated measures methods. Regretfully, multilevel models have yet to make significant headway into sport biomechanics despite being present in other sport performance domains, such as sport psychology [2, 3, 6]. If fact, to our knowledge, the use of multilevel modeling in

---

[i]ANOVA search: ["ANOVA" or "repeated measures" source:biomechanics], multilevel modeling search: ["multilevel regression" OR "mixed-effects" OR "mixed-model" source:biomechancis]. Date range limited between 2012 and 2022. Disregarding patents and citations. Performed February 2022.

sport biomechanics is limited to three papers, all published since 2019 [13, 26, 27]. While studying the baseball pitching motion Slowik et. al. [26] and Solomito et al [27] used a multilevel framework to examine the associations between various kinematic factors and pitching arm joint loads. Additionally, Iglesias et. al. examined between-participant differences in the within-participant load-velocity relationship during several weightlifting exercises [13]. Although these papers provide important insight into their respective research areas, we feel a formal introduction of the advantages, disadvantages, and limitations of multilevel modeling would benefit those in sport biomechanics working alongside athletes in longitudinal settings and who frequently encounter repeated measures study designs. Therefore, our purpose is to introduce multilevel regression modeling to a sports biomechanics audience and offer a brief tutorial on model notation, construction, and interpretation. We also outline the advantages and disadvantages of multilevel modeling over traditional repeated measures designs and offer recommendations for those interested in further exploration of multilevel techniques.

## 2  The Traditional Regression Model

Before examining the multilevel regression model, let us first consider a simple linear regression predicting some outcome measure, $y$, from one predictor variable, $x$, given by Equation 1:

$$\hat{y}_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{1}$$

In Equation 1, $\beta_0$ represents the model intercept and $\beta_1$ represents the model slope. Recall that a model intercept can be thought of as the predicted value of $y$ when $x$ equals zero and a model slope can be thought of as the predicted increase (or decrease, if negative) in $y$ for a 1-unit increase in $x$. If our predictor, $x$, is centered around its mean, then the intercept represents the predicted value of $y$ for an average value of $x$. $\epsilon_i$ represents the residual error between the $i$-th model prediction and the $i$-th observation $(y_i - \hat{y}_i)$ (Figure 1). The one-predictor linear regression model can be extended to multiple regression by adding additional predictor variables (i.e. $x_2$, $x_3...x_n$) or nonlinear[ii] regression by adding polynomial (i.e. $x^2, x^3...x^n$), exponential $(e^x)$, or logarithmic $(ln(x))$ terms. Additionally, how predictor variables' effects depend on one another can be examined by introducing their product terms (i.e. $x_1 * x_2$) and analyzing their interactions.

---

[ii]More formal writings will use *linear vs. nonlinear* in terms of the role of the regression coefficients (i.e. $\beta_0$, $\beta_1$, etc...). However, here we use *nonlinear* in terms of the relationships between $x$ and $y$. For more details we refer the reader to the works of Curran-Everett [7]
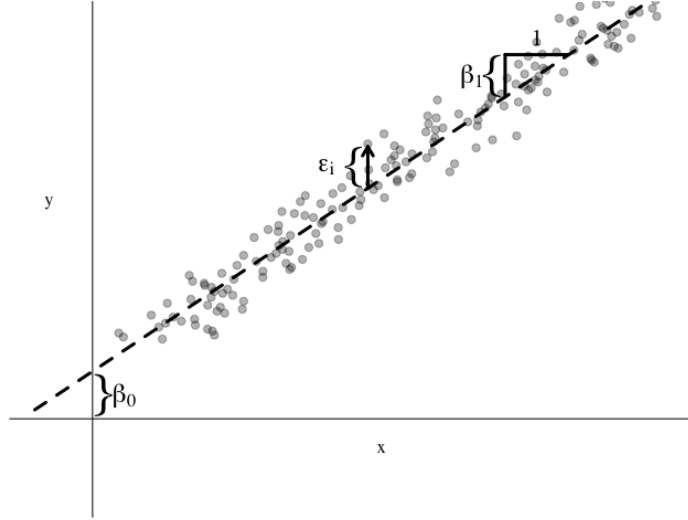
Figure 1: Hypothetical Simple Regression

## 3  Nested Data and the Multilevel Regression Model for Repeated Measures

Collecting multiple observations from a single participant results in what is sometimes referred to in other fields as 'nested', 'hierarchical', or 'multilevel' data [1, 4, 23, 25, 29]. In repeated measures designs, we consider observations to be nested within individuals. In other applications, individuals may be nested within some grouping structure such as classrooms, congressional districts, or treatment arms. Regardless of the grouping structure in question, nested data violate the assumption of observational independence required by traditional regression [8, 12, 23]. That is, two observations from the same individual will be more closely related to each other compared to two observations from two different individuals.

The degree to which nested data are more closely related than independent data is sometimes referred to as the *intraclass correlation* and is computed as the variance in a model's outcome accounted for by the data's nested structure divided by the outcome's total variance [12, p. 15][iii].

$$ICC = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2} \tag{2}$$

Where $\sigma_{u_0}^2$ and $\sigma_e^2$ represent the between-group and within-group variances, respectively. Under highly dependent data conditions (where most variance is accounted for by the nested data structure), the ICC will approach 1.0 since Equation 2 simplifies to $\frac{\sigma_{u_0}^2}{\sigma_{u_0}^2}$ as $\sigma_e^2$ trends towards zero. Conversely, the ICC will approach

---

[iii]For those more familiar with the ANOVA-based framework, an analog to the ICC is the between sum of squares divided by the total sum of squares ($\frac{SS_{btwn}}{SS_{tot}}$).

3

zero under highly independent data since Equation 2 simplifies to $\frac{0}{\sigma_e^2}$ as $\sigma_{u_0}^2$ trends towards zero.

In practice, the ICC will lay somewhere between 0 and 1. Most important for the applied researcher, even apparently small deviations from 0 can drastically inflate type 1 error rates if ordinary least squares methods are used [18].

---

Similar to traditional regression, we can start with a simple multilevel model predicting some outcome measure, $y$, from one predictor variable, $x$. For this paper we will be following Hox's [12] notation[iv].

$$\hat{y}_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \epsilon_{ij} \tag{3}$$

Equation 3 is often referred to as the *level-one* or *first-level* model since it deals with the first level of nesting in our data.

Equation 3 is identical to Equation 1 except we now have a second subscript, $j$, denoting the nested nature of the data. $x_{ij}$ now represents observation $i$ from individual $j$ while $\beta_{0j}$ and $\beta_{1j}$ represent the model intercept and slope for individual $j$. $\hat{y}_{ij}$ represents the model prediction for observation $i$ from individual $j$ and $\epsilon_{ij}$ represents the residual error between each observed outcome and its corresponding predicted value $(y_{ij} - \hat{y}_{ij})$. The inclusion of a second subscript, $j$, thereby allowing model intercepts and slopes to vary between individuals is the fundamental difference between traditional and multilevel regression.

By allowing model intercepts and slopes to vary between individuals, the multilevel model for repeated measures parses each individual's intercept and slope into a combination of the average model intercept/slope plus some individual-specific deviation from the average. Equations 4 and 5 represent the regression equations predicting each individual's model intercept (Equation 4) and slope (Equation 5) and make up the second-level model. In Equations 4 and 5, the first term on the right side ($\gamma_{00}$ and $\gamma_{10}$) represent the *average* model intercept and slope, respectively, across all individuals. These average model parameters are known as *fixed*-effects since they do not vary between individuals (notice there is no $j$ subscript). The second terms on the right side ($\mu_{0j}$ and $\mu_{1j}$) represent the individual-specific deviations from the average model intercept and slope, respectively. Because $\mu_{0j}$ and $\mu_{1j}$ vary between individuals, they are known as *random*-effects. Modeling parameter variability as a combination of fixed and random-effects is why multilevel models are sometimes referred to as *mixed*-effects models (a mixture of fixed and random effects).

[iv]It should be known that other authors have used slightly (or sometimes very) different notations. Admittedly, this is one of the higher barriers to entry for those without strong analytical training. We encourage readers to take considerable time learning model notation as it will accelerate the remaining stages of the learning process

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \tag{4}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \tag{5}$$

The multilevel model can then use second-level parameters $(z_j)$ to account for variance in the predicted intercepts and slopes. With the inclusion of second-level parameters, $\mu_{0j}$ and $\mu_{1j}$ represent the residual variation in the predicted intercept and slope after accounting for $z_j$.

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + \mu_{0j} \tag{6}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_j + \mu_{1j} \tag{7}$$

Substituting the right side of Equations 6 and 7 for $\beta_{0j}$ and $\beta_{1j}$ from Equation 3 yields the complete multilevel model predicting one outcome from one first-level and one second-level variable:

$$\hat{y}_{ij} = \overbrace{\gamma_{00} + \gamma_{01}z_j + \mu_{0j}}^{intercept} + \overbrace{(\gamma_{10} + \gamma_{11}z_j + \mu_{1j})}^{slope} x_{ij} + \epsilon_{ij} \tag{8}$$

Distributing $x_{ij}$ and arranging like terms helps delineate the fixed and random parts of the model:

$$\hat{y}_{ij} = \overbrace{\gamma_{00} + \gamma_{01}z_j + \gamma_{10}x_{ij} + \gamma_{11}x_{ij}z_j}^{fixed} + \overbrace{\mu_{1j}x_{ij} + \mu_{0j} + \epsilon_{ij}}^{random} \tag{9}$$

As with traditional regression models, the multilevel regression model can be generalized to multiple multilevel[v] regression with $m$ first-level and $n$ second-level predictors:

$$\hat{y}_{ij} = \beta_{0j} + \sum_{m=1}^{m} \beta_{mj}x_{mij} + \epsilon_{ij} \tag{10}$$

$$\beta_{0j} = \gamma_{00} + \sum_{n=1}^{n} \gamma_{0n}z_{nj} + \mu_{0j} \tag{11}$$

$$\beta_{mj} = \gamma_{m0} + \sum_{n=1}^{n} \gamma_{1n}z_{nj} + \mu_{nj} \tag{12}$$

---

[v]yeah, "multiple multilevel", I know. Yikes

# 4 Benefits and Drawbacks of Multilevel Models over Traditional Repeated Measures ANOVA

The three assumptions made by repeated measures ANOVA that are most relevant to sport biomechanics: temporal equivalence, design balance, and case completeness.

|  | Repeated Measures ANOVA | Multilevel Regression |
|---|---|---|
| Temporal equivalence[a] | Assumes repeated measures are equally spaced in time | Repeated measures may be unequally spaced in time |
| Balanced design | Assumes participants have same number of repeated measures | Participants may have different number of repeated measures |
| Case completeness | Assumes no missing data; incomplete cases are handled through listwise deletion | Missing data may be accounted for under certain circumstances[b] |

[a] "Temporal" equivalence does not only apply to growth models, where time is placed on the x-axis. It may be extended to "whatever is on the x-axis of your scatterplot".
[b] Missing data analysis is an extremely rich stand-alone area of research that we cannot do justice in this article alone. Readers are directed to works of Enders [9] for a more thorough exploration of the topic

## 4.1 When ANOVA will do just fine

Simple pre-post designs with no missing data
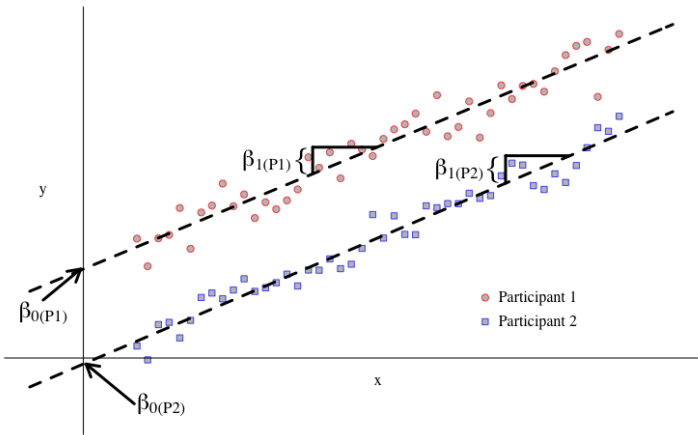
## 5 Other Figures to incorporate



Figure 2: Fictional multilevel model with random intercepts. Individual regression lines are parallel (equal slopes) but differ in their respective intercepts
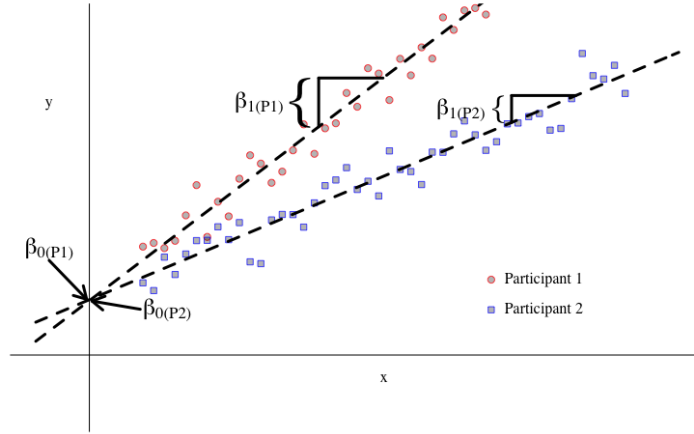
Figure 3: Fictional multilevel model with random slopes. Individual regression lines are not parallel (unequal slopes) but share a mutual intercept
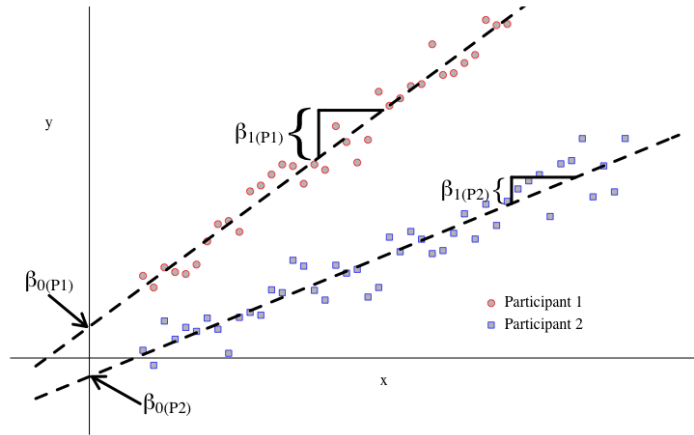


Figure 4: Fictional multilevel model with random intercepts and slopes. Individual regression lines are not parallel and do not share a mutual intercept

## 6   Discard Pile

We biomechanists like to throw away lots of data. Our cameras capture the positions of dozens of reflective markers with sub-millimeter accuracy hundreds of times each second while our participants perform several, sometimes dozens, of movement trials during data collection. Each new trial brings additional data that could increase statistical power and be used to better understand our research questions. Yet, far too often, choices related to experimental design or statistical analysis greatly reduce the volume of data we work

with. Reduced data volume negatively impacts inferential power when conducting statistical analyses and can potentially indicate inefficient use of data collection resources and research labor.

Two research practices common in sports biomechanics that reduce data volume are discretization and ensemble averaging. Discretization occurs when researchers extract one value, such as a maxima, minima, or average value, from a continuous time series of data. Ensemble averaging occurs when researchers combine several trials from the same participant into one *representative* trial. Discretization and ensemble averaging can also be combined to further reduce the available data by averaging discrete values from multiple trials or taking discrete values from ensemble averaged time series. Although sometimes appropriate depending on the research question, discretization and ensemble averaging reduce the dimensionality and variability in time series data, reducing statistical power.

While discretization can be ameliorated though emerging techniques such as statistical parametric mapping [22], researchers may still wish to examine certain discrete measures if warranted by their domain expertise and research question of interest. One such scenario is the repeated measure of biomechanic or performance values over time. Typically, repeated measures data in sports biomechanics are examined using univariate or multivariate repeated measures analysis of (co)variance. While these statistical tools are appropriate under certain conditions, they make several assumptions that may often be problematic in the observational or quasi-experimental settings common in sport biomechanics research. In such cases, multilevel regression modeling may complement or replace more traditional repeated measures analyses.

Although multilevel techniques are present in other sport performance domains such as sport psychology [2, 3, 6], they have yet to make significant headway into sport biomechanics. If fact, to our knowledge, the use of multilevel modeling in sport biomechanics is limited to two papers, both published since 2019 [13, 26]. Slowik et. al. used a multilevel framework to contrast the strong within and weak between-participant relationships between elbow joint loading and baseball pitching speed [26] while Iglesias et. al. examined between-participant differences in the within-participant load-velocity relationship during several weightlifting exercises [13]. Although these two papers provide important insight into their respective research areas, we feel a formal introduction of the advantages, disadvantages, and limitations of multilevel modeling would benefit those in sport biomechanics working with repeated measures designs. Therefore, our purpose is to introduce multilevel regression modeling to a sports biomechanics audience and offer a brief tutorial on model notation, construction, and interpretation. We also outline the advantages and disadvantages of multilevel modeling over traditional repeated measures designs and offer recommendations for those interested in further exploration of multilevel techniques.

Some of the more notable assumptions that can cause problems for sport biomechanists dealing with

repeated measures include temporal equivalence (i.e. time between data points must be the same for all participants), balanced study designs (i.e. must have the same number of data points for each individual) and case completeness (i.e. must have no missing data).

Compared with traditional univariate or multivariate analysis of (co)variance, multilevel modeling allows more researcher flexibility through fewer statistical assumptions and the ability to handle missing data, non-linear relationships, and time-varying covariates [12].

In addition to its assumptions regarding data symmetry and structure, RM·ANOVA is restricted to testing for the presence of linear relationships between variables and, while the ANOVA framework can handle time-invariant covariates, modeling covariates which change over the duration of a longitudinal study is not possible in RM·ANOVA.

## References

[1] Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., and Van Der Sluis, S. A solution to dependency: using multilevel analysis to accommodate nested data. *Nature neuroscience*, 17(4):491–496, 2014.

[2] Beauchamp, M. R., Bray, S. R., Fielding, A., and Eys, M. A. A multilevel investigation of the relationship between role ambiguity and role efficacy in sport. *Psychology of Sport and Exercise*, 6(3):289–302, 2005.

[3] Benson, A. J., Šiška, P., Eys, M., Priklerova, S., and Slepička, P. A prospective multilevel examination of the relationship between cohesion and team performance in elite youth sport. *Psychology of Sport and Exercise*, 27:39–46, 2016.

[4] Boisgontier, M. P. and Cheval, B. The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews*, 68:1004–1005, 2016.

[5] Casals, M. and Finch, C. F. Sports biostatistician: a critical member of all sports science and medicine teams for injury prevention. *Injury prevention*, 23(6):423–427, 2017.

[6] Cornelius, A. E., Brewer, B. W., and Van Raalte, J. L. Applications of multilevel modeling in sport injury rehabilitation research. *International Journal of Sport and Exercise Psychology*, 5(4):387–405, 2007.

[7] Curran-Everett, D. Explorations in statistics: regression. *Advances in Physiology Education*, 35(4):347–352, 2011.

[8] DeHart, W. B. and Kaplan, B. A. Applying mixed-effects modeling to single-subject designs: An introduction. *Journal of the Experimental Analysis of Behavior*, 111(2):192–206, 2019.

[9] Enders, C. K. *Applied missing data analysis.* Guilford press, 2010.

[10] Glazier, P. S. and Mehdizadeh, S. Challenging conventional paradigms in applied sports biomechanics research. *Sports Medicine*, 49(2):171–176, 2019.

[11] Halperin, I., Vigotsky, A. D., Foster, C., and Pyne, D. B. Strengthening the practice of exercise and sport-science research. *International Journal of Sports Physiology & Performance*, 13(2), 2018.

[12] Hox, J. J., Moerbeek, M., and Van de Schoot, R. *Multilevel analysis: Techniques and applications.* Routledge, 2017.

[13] Iglesias-Soler, E., Mayo, X., Rial-Vázquez, J., and Haff, G. G. Inter-individual variability in the load-velocity relationship is detected by multilevel mixed regression models. *Sports biomechanics*, 20(3):304–318, 2021.

[14] Knudson, D. Significant and meaningful effects in sports biomechanics research. *Sports Biomechanics*, 8(1):96–104, 2009.

[15] Knudson, D. Confidence crisis of results in biomechanics research. *Sports biomechanics*, 16(4):425–433, 2017.

[16] Liao, T.-C., Pedoia, V., Majumdar, S., and Souza, R. B. Longitudinal changes of patellar alignment before and after anterior cruciate ligament reconstruction with hamstring autograft. *The American Journal of Sports Medicine*, 49(11):2908–2915, 2021.

[17] Mueske, N. M., Patel, A. R., Pace, J. L., Zaslow, T. L., VandenBerg, C. D., Katzel, M. J., Edison, B. R., and Wren, T. A. Improvements in landing biomechanics following anterior cruciate ligament reconstruction in adolescent athletes. *Sports biomechanics*, 2018.

[18] Musca, S. C., Kamiejski, R., Nugier, A., Méot, A., Er-Rafiy, A., and Brauer, M. Data with hierarchical structure: impact of intraclass correlation and sample size on type-i error. *Frontiers in Psychology*, 2:74, 2011.

[19] Nielsen, R. Ø., Chapman, C. M., Louis, W. R., Stovitz, S. D., Mansournia, M. A., Windt, J., Møller, M., Parner, E. T., Hulme, A., Bertelsen, M. L., et al. Seven sins when interpreting statistics in sports injury science, 2018.

[20] Nielsen, R. Ø., Shrier, I., Casals, M., Nettel-Aguirre, A., Møller, M., Bolling, C., Bittencourt, N. F., Clarsen, B., Wedderkopp, N., Soligard, T., et al. Statement on methods in sport injury research from the first methods matter meeting, copenhagen, 2019. *Journal of Orthopaedic & Sports Physical Therapy*, 50(5):226–233, 2020.

[21] Nojiri, S., Yagi, M., Mizukami, Y., and Ichihashi, N. Static stretching time required to reduce iliacus muscle stiffness. *Sports biomechanics*, 2019.

[22] Pataky, T. C., Vanrenterghem, J., Robinson, M. A., and Liebl, D. On the validity of statistical parametric mapping for nonuniformly and heterogeneously smooth one-dimensional biomechanical data. *Journal of biomechanics*, 91:114–123, 2019.

[23] Raudenbush, S. W. and Bryk, A. S. *Hierarchical linear models: Applications and data analysis methods*, volume 1. sage, 2002.

[24] Sainani, K. L., Borg, D. N., Caldwell, A. R., Butson, M. L., Tenan, M. S., Vickers, A. J., Vigotsky, A. D., Warmenhoven, J., Nguyen, R., Lohse, K. R., et al. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. *British Journal of Sports Medicine*, 55(2):118–122, 2020.

[25] Schielzeth, H. and Nakagawa, S. Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution*, 4(1):14–24, 2013.

[26] Slowik, J. S., Aune, K. T., Diffendaffer, A. Z., Cain, E. L., Dugas, J. R., and Fleisig, G. S. Fastball velocity and elbow-varus torque in professional baseball pitchers. *Journal of athletic training*, 54(3):296–301, 2019.

[27] Solomito, M. J., Garibay, E. J., Cohen, A. D., and Nissen, C. W. The influence of trunk mechanics on the assessment of elbow position in college aged pitchers. *Sports Biomechanics*, pages 1–11, 2020.

[28] Stronska, K., Golas, A., Wilk, M., Zajac, A., Maszczyk, A., and Stastny, P. The effect of targeted resistance training on bench press performance and the alternation of prime mover muscle activation patterns. *Sports Biomechanics*, pages 1–15, 2020.

[29] Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., and Smith, G. M. Mixed effects modelling for nested data. In *Mixed effects models and extensions in ecology with R*, pages 101–142. Springer, 2009.