

Sample size justifications in Gait & Posture

Citation for published version (APA):

McCrum, C., van Beek, J., Schumacher, C., Janssen, S., & Van Hooren, B. (2022). Sample size justifications in Gait & Posture. *Gait & Posture*, 92, 333-337. <https://doi.org/10.1016/j.gaitpost.2021.12.010>

Document status and date:

Published: 01/02/2022

DOI:

[10.1016/j.gaitpost.2021.12.010](https://doi.org/10.1016/j.gaitpost.2021.12.010)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

CC BY

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

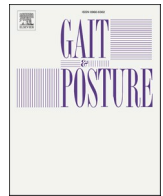
www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Sample size justifications in Gait & Posture

Christopher McCrum^{*}, Jorg van Beek, Charlotte Schumacher, Sanne Janssen, Bas Van Hooren

Department of Nutrition and Movement Sciences, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre+, Maastricht, The Netherlands

ARTICLE INFO

Keywords:
Metascience
Power analysis
Statistical power
Study design
Resources

ABSTRACT

Background: Context regarding how researchers determine the sample size of their experiments is important for interpreting the results and determining their value and meaning. Between 2018 and 2019, the journal Gait & Posture introduced a requirement for sample size justification in their author guidelines.

Research Question: How frequently and in what ways are sample sizes justified in Gait & Posture research articles and was the inclusion of a guideline requiring sample size justification associated with a change in practice?

Methods: The guideline was not in place prior to May 2018 and was in place from 25th July 2019. All articles in the three most recent volumes of the journal (84–86) and the three most recent, pre-guideline volumes (60–62) at time of preregistration were included in this analysis. This provided an initial sample of 324 articles (176 pre-guideline and 148 post-guideline). Articles were screened by two authors to extract author data, article metadata and sample size justification data. Specifically, screeners identified if (yes or no) and how sample sizes were justified. Six potential justification types (Measure Entire Population, Resource Constraints, Accuracy, *A priori* Power Analysis, Heuristics, No Justification) and an additional option of Other/Unsure/Unclear were used.

Results: In most cases, authors of Gait & Posture articles did not provide a justification for their study's sample size. The inclusion of the guideline was associated with a modest increase in the percentage of articles providing a justification (16.6–28.1%). *A priori* power calculations were the dominant type of justification, but many were not reported in enough detail to allow replication.

Significance: Gait & Posture researchers should be more transparent in how they determine their sample sizes and carefully consider if they are suitable. Editors and journals may consider adding a similar guideline as a low-resource way to improve sample size justification reporting.

1. Introduction

Context regarding the sample size determination of an experiment is important for interpreting and determining the meaning of the experiment's results. For this reason, many reporting guidelines, such as the CONSORT Statement and extensions [1–5], the STROBE statement [6] and the recently published PRESENT 2020 checklist (Proper Reporting of Evidence in Sport and Exercise Nutrition Trials) [7], explicitly require reporting and justification of sample size calculations. Despite this, a lack of sample size justification is frequently reported as a source of bias by systematic reviews across fields. Lately, analysis of the reporting of sample size justification has become a topic of interest in many fields e.g. [8–10]. Between 2018 and 2019, the journal Gait & Posture introduced the following statement in their author guidelines:

“All Articles should include a justification of their sample size. While there is no set requirement for minimum sample size, studies considered to

have too small a sample size to answer the research question will be rejected.” [11].

This guideline was introduced in parallel with an editorial on “Statistical guidance for a successful research paper” [12] by the journal's Statistical Editor, supported by the journal editors [13]. As the only journal exclusively focusing on gait and posture research and as the official journal of the International Society of Posture and Gait Research (ISPGR), we interpret these publications and guideline inclusion as a positive step for our field, indicating a willingness to improve how we plan and analyse our studies.

There has not yet been a systematic analysis of sample size justifications in the field of gait and posture research. It is therefore unclear how researchers in this field determine their sample sizes and how these determinations are reported. There are many possible justifications of a study's sample size and each has its own implications for both study design and interpretation of results [14]. It would, for example, be

^{*} Correspondence to: Department of Nutrition and Movement Sciences, Maastricht University, PO Box 616, Maastricht 6200MD, The Netherlands.
E-mail address: chris.mccrum@maastrichtuniversity.nl (C. McCrum).

valuable for the field to determine whether sample sizes are generally justified using *a priori* power analyses based on either an expected effect size or minimum effect size of interest and if these effect sizes are well established or not. Alternatively, limited resources may be a larger driving force for the sample sizes seen in this field, particularly due to its overlap with clinical and athletic populations (limited number of available participants), indicating a need for large collaborative projects, or which might manifest as financial limitations associated with equipment use, which might again indicate the need for large collaborations or further development of lower cost equipment. The recent addition of this guideline also provides an opportunity to examine if this relatively simple change is associated with a positive change in practice with more transparent declaration of how sample sizes were determined. Once transparent reporting becomes the norm, it will create opportunities for more rigorous field-wide analyses of these issues.

To make an initial evaluation of sample size justifications in this field, we screened samples of articles from *Gait & Posture*, including both the most recent volumes, and the latest volumes before the inclusion of this guideline. We also determined the frequency of various types of sample size justifications based on the most common justification types outlined in “Sample Size Justifications” by Lakens [14] (Table 1).

2. Method

This study’s rationale, data collection and analysis plans were pre-registered on the Open Science Framework [15] and all materials and data can be found on the OSF Project Page [16].

2.1. Determining the date of introduction of the author guideline

Following communication with colleagues, checking the website “WayBack Machine” (<https://web.archive.org/>) and web searches for PDF versions of the author guidelines, it was determined that the guideline was not included in the May 2018 version of the author guidelines (copies of the guidelines were found in the appendices of two postgraduate dissertations from Universidade de Brasília [17,18]) and was included in the September 2019 version. Following communication with the Deputy Editor of *Gait & Posture*, it was determined that while a precise date of the inclusion of the guideline was not known, the publisher confirmed that it was included in a version of the guidelines dated 25th July 2019 (*J. Stebbins, personal communication*). As a result, we could confirm that the guideline was not in place prior to May 2018 and was in place from 25th July 2019.

2.2. Sample size determination

This was a short-term project with time limitations. Therefore, a

Table 1
Overview of Possible Justifications for the Sample Size in a Study (adapted from Lakens [14]).

Type of Justification	When is this Justification Applicable?
Measure entire population	<i>The researcher can specify the entire population, it is finite, and it is possible to measure (almost) every entity in the population.</i>
Resource constraints	<i>Limited resources are the primary reason for the choice of the sample size the researcher can collect.</i>
Accuracy	<i>The research question focusses on the size of a parameter, and a researcher collects sufficient data to have an estimate with a desired level of accuracy.</i>
A-priori power analysis	<i>The research question has the aim to test whether certain effect sizes can be statistically rejected with a desired statistical power.</i>
Heuristics	<i>The researcher decides upon the sample size based on a heuristic, general rule or norm that is described in the literature or communicated orally.</i>
No justification	<i>The researcher has no reason to choose a specific sample size, or does not have a clearly specified inferential goal and wants to communicate this honestly</i>

sample size of articles to screen was determined based on this resource constraint. Given the weekly time commitment feasible for the authors and the period within which the data collection should be completed, as well as an estimated rate of 13 articles screened per hour based on piloting (see pilot details below), a total of approximately 312 articles, each screened by two authors, was considered feasible. Since the earliest confirmed date of guideline inclusion (25th July 2019) until and including May 2021, 14 complete volumes of the journal were published (Volume 73–86) including approximately 600 articles (reviews excluded). In order to have a sample most representative of current practice, the three most recent volumes of the journal (84–86) and the three most recent, pre-guideline volumes (60–62) at time of preregistration were included. This provided an initial sample of 324 articles (176 pre-guideline and 148 post-guideline).

2.3. Procedures

Prior to data collection and screening, the authors read and discussed the article “*Sample Size Justification*” by Lakens [14]. The purpose was to ensure that all authors were aware of the concepts and issues in this topic and would understand the characteristics of common justifications of sample size and how to distinguish them from each other.

To standardise the screening process, a draft instruction sheet and standardised excel results sheet were prepared to guide the screeners. Items included *Screener Initials, Article First Author Surname, Title, Publication Year, Journal Volume, DOI number, Received Date, Accepted Date, If the article involved live human participants* (no secondary analyses or retrospective database studies were considered) and *Total Sample Size*. Additionally, screeners were asked if a sample size justification was provided (yes or no) and if so, in which section of the article it was found and what type of justification was used. Six potential justification types were specifically included as options based on the article of Lakens [14] (Table 1) and an additional option of Other/Unsure/Unclear was included, with the intention that this would be discussed after screening was complete. Note that a “No Justification” justification was only selected when the authors explicitly stated that they had no justification for their chosen sample size. Screeners were also asked to copy and paste the relevant text from the articles into the excel sheet.

A pilot data collection on eight preselected articles was conducted to test the draft screening instructions and to have an estimate of the time required for article screening. The project leader (CM) pre-selected eight articles to include a range of justification types, detail and perceived clarity. The other four authors then performed the screening and timed their work. On average, it took approximately 4.5 min per article (mean=4.56 min, SD=0.98 min). Some contradictions in responses and classifications were found and these were discussed. The screening instructions were then updated based on identified issues and to reduce the risk of human error (i.e. overlooking relevant text). The draft and final screening instructions, as well as the pilot results file can be found on the OSF Project Page [16]. The final screening instructions are included in the [supplementary materials](#).

In preparation for the full article screening, the lead author (CM) downloaded all non-review articles per volume from the journal website in PDF format. Articles were combined from the target volumes in a single folder and arranged alphabetically with the title as filename. Based on the availability of the various authors, these articles were distributed among the authors such that all articles would be screened by two authors. Following screening, the lead author combined the results in a single sheet and identified disagreements. These disagreements were resolved by a combination of the original screeners re-evaluating their assessment, discussion between the screeners and a third author acting as a tie breaker. It was planned *a priori* to report the results in a descriptive manner in terms of frequencies of the outcomes relative to the pre- and post-guideline samples [15].

3. Results

324 articles (176 pre-guideline and 148 post-guideline) from the non-reviews section of Gait & Posture volumes 60–62 and 84–86 were screened. Of these, one corrigendum and one author response article were excluded. Articles reporting secondary analysis of data not involving live participants for the first time were then identified and excluded (17 pre-guideline articles and 20 post-guideline articles). The presence and type of sample size justifications were determined for the remaining 157 pre-guideline and 128 post-guideline articles.

Pre-guideline, 26 articles (16.6%) included an identifiable justification of their sample sizes and post-guideline, 36 articles (28.1%) included a justification (Fig. 1). *A priori* power analyses were the dominant type of justification (Table 2). The proportion of articles with no justification decreased post-guideline and the proportion of articles where the sample size justification was unclear slightly increased (Fig. 1).

Eleven pre-guideline articles and 14 post-guideline articles were considered to have unclear sample size justifications. Pre-guideline, the 11 articles involved the authors referring to convenience samples, but without enough detail to determine whether the sample size itself (as opposed to the method of recruitment) was chosen out of convenience, which we could have considered a resource-based justification. Post-guideline, the 14 articles with unclear sample size justifications included 11 mentioning convenience samples but lacking other details. Two of the 14 studies claimed that sample sizes were based on previous research but did not provide sufficient detail to determine if this led to a power calculation or a heuristics-type justification and one study justified their sample size with a *post hoc* sensitivity power analysis that was clearly not used to determine the sample size *a priori*.

We originally planned to describe and not analyse the sample sizes of the study, but it is interesting to note that the median sample size in each category of justification decreased over time, particularly in the articles with sample size justifications (Fig. 2). Sample sizes of the included articles, categorised by time point and justification status, are shown in Fig. 2. Overall, the median sample size was 30, with median sample sizes of 29.5, 30 and 30 for articles with, without and with unclear sample size justifications, respectively.

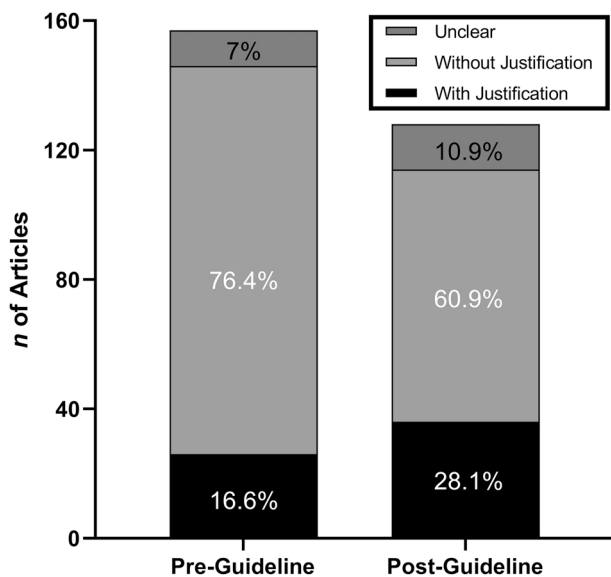


Fig. 1. Number of articles screened pre- (Gait & Posture Volumes 60–62) and post-guideline (Gait & Posture Volumes 84–86) with percentages of articles with and without justifications and with unclear justifications.

Table 2
Types of Sample Size Justifications Reported Pre- and Post-Guideline.

	Pre-Guideline	Post-Guideline
Measure entire population	0	0
Resource constraints	1	0
Accuracy	0	0
A-priori power analysis	23	34
Heuristics	2	0
No justification	0	0
Other/Combination	0	2

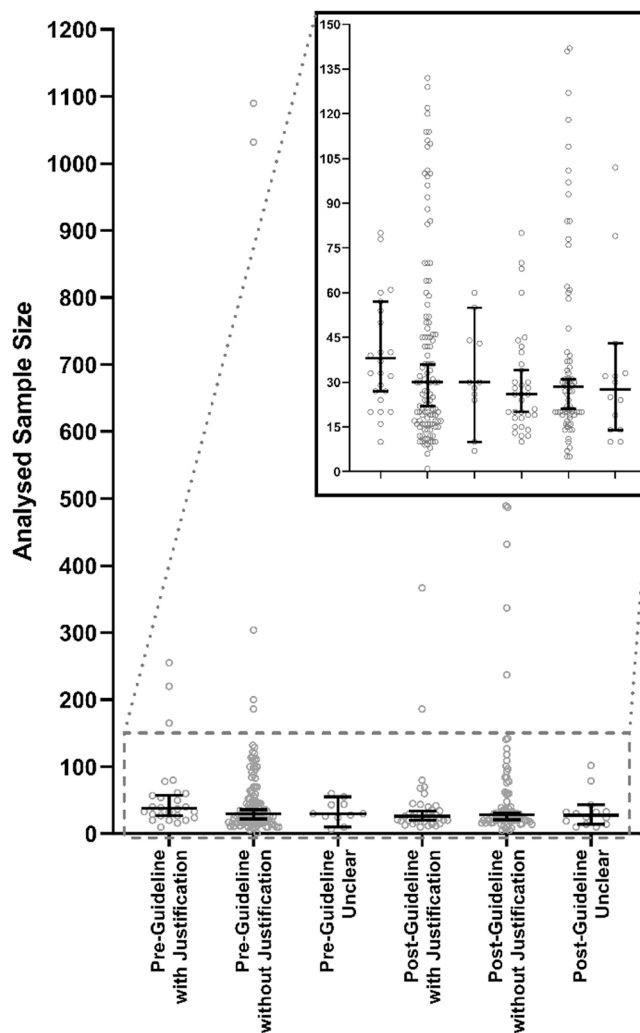


Fig. 2. Analysed sample sizes of the included studies categorised by publication pre- (Gait & Posture Volumes 60–62) and post-guideline (Gait & Posture Volumes 84–86) and by the presence of sample size justification or lack thereof (individual data points, medians and 95% confidence intervals).

4. Discussion

We aimed to evaluate sample size justifications in the gait and posture research field and to make an initial assessment of whether the inclusion of a sample size justification requirement in a journal’s author guidelines would be associated with a positive change in practice. Our results from 157 pre-guideline and 128 post-guideline articles demonstrate that authors of Gait & Posture articles do not provide a justification for their study’s sample size in the majority of cases and that the inclusion of the author guideline requiring such a justification was associated with a modest increase in the percentage of articles providing

a justification (16.6–28.1%).

Given that only 28.1% of post-guideline articles include a sample size justification, we conclude that adherence to this guideline is not a (consistent) criteria for publication enforced by the journal or publisher. Nevertheless, it does appear that the inclusion of the guideline is associated with an, albeit modest, increase in sample size justification reporting. As a result, it seems that simply including a few sentences in a journal's author guidelines may be sufficient to have a small positive influence on practice and can certainly be recommended for other journals looking for low-resource ways in which to improve the transparency of articles published. However, due to the nature of the study, we are not able to determine causality between the inclusion of the guideline and the changes in sample size justification and can only infer an association. One other reason for the observed changes may be an increased awareness among researchers of the importance of explicitly providing sample size justifications and increased openness and transparency more generally. There have been related publications and developments in adjacent and overlapping fields of sports science, biomechanics and physiology [7,19–24] that may have played a role in this way. One other reason may be changes over time in funder, university or ethical review board criteria that filter through to the published manuscripts. It should also be kept in mind that change in the proportion of articles reporting sample size justifications is no guarantee that the justifications are correct or appropriate. Rigorous sample size planning, justification and power analyses can be time consuming and require considerable expertise. Requiring justification without any quality control may lead to an increase in poorly or nefariously conducted power analyses (or "power analysis hacking") [25] simply in order to fulfil this requirement. Asking authors to produce well planned and substantiated justifications for their sample sizes at the journal submission stage may be too late to ensure quality since the study has already been conducted. We strongly recommend that researchers spend time prior to data collection considering and planning what sample size will be appropriate to address their aims and that ethical review boards place more focus on this aspect of study design, while adjustments can still be made.

As can be seen in our data (available in the [supplementary material](#) and on the OSF Project Page [16]), many of the *a priori* power analyses performed in the included studies based their expected effect sizes on single (often pilot) studies and smallest effect sizes of interest or clinically meaningful differences were rarely mentioned. Pilot studies serve particular purposes in the research process, but suitably powered hypothesis testing is not typically one of them and as such, these small, underpowered analyses are not a reliable basis for power calculations [14,26–28]. Combined with a lack of detail on smallest effect sizes of interest or clinically meaningful effects sizes, this creates difficulty drawing or trusting the conclusions of a study; it is often not clear if the effect size powered for is reasonable or meaningful [29]. Additionally, the articles screened often did not provide sufficient information to replicate the power analyses. For example, often the values and justifications used to determine the effect sizes and the type of statistical test used were not specified. Each of these issues potentially reduces the quality, reliability and interpretability of the studies [30]. Similar results have been found in other fields [9,10,21,31–36]. For example, in orthopaedics, it has been reported that the majority of studies with power analyses do not provide rationale for the effect size used in their power calculations [9] and that almost half of the power analyses cannot be replicated [35]. In the Journal of Biomechanics in 2018–2019, 29 of 653 experimental papers reported conducting an *a priori* power analysis, about half of which could be replicated using the information provided [36]. While these issues are not unique to our field, power analyses urgently require attention in gait and posture research and the data provided in this article may provide a starting point.

It is worth noting that only one study justified their sample size based on resource constraints. However, most research is affected by resource constraints to some extent, whether that be lab time, contract length,

cost of measurements or available funding, or more recently, pandemic-related lab closures. As Lakens [14] highlights, while resource constraints may not be the primary justification of sample sizes, they are usually secondary justifications. We therefore recommend that researchers consider and report these restrictions when justifying their sample sizes in the future and we refer readers to previous work discussing this issue [14,37].

While not the focus of this analysis, we did not observe many differences in sample size between articles with and without justifications and overall, we found a median sample size of 30 participants. This is slightly higher than the median sample size in other related fields, for example in a sports science journal from 2017 to 2020 (median of 19 participants) [21] and in three applied biomechanics journals in 2009 (median range of 12–18 participants) [19,38]. While larger, 30 participants may not be sufficient to detect meaningful changes in some parameters. For example, very small changes (in absolute terms) in measures of step variability may be clinically meaningful [39] and could require anywhere between 39 and 192 participants (depending on the correlation between repeated measures) in a within participant (i.e. pre-post) comparison [40]. Similar to our findings in Gait & Posture, it has been reported that sample sizes in the Journal of Applied Biomechanics and the ISBS (International Society of Biomechanics in Sports) Proceedings did not appreciably change over time (1989–2009) [19,41].

It should be kept in mind that we analysed three issues pre-guideline and three issues post-guideline. As Gait & Posture publishes eight regular issues per year, we feel that this will be a representative sample of the articles published by the journal, but the decision to include six issues was based on resources constraints and not on a sample size calculation based on power or precision of analysis. Also, the results from this analysis are specific to one journal and while this is the single society-associated journal for research on posture and gait, it is by no means the only publication outlet for this community and research field.

5. Conclusion

In the majority of cases, authors of Gait & Posture articles do not provide a justification for their study's sample size. The inclusion of an author guideline requiring such a justification was associated with a modest increase (16.6–28.1%) in the percentage of articles providing a justification from volumes 60–62 (February, March and May 2018) to volumes 84–86 (February, March and May 2021). However, it is important to note that encouraging reporting of sample size justifications does not ensure that these justifications are appropriate or of high quality (the majority of the observed power analyses were not reported in sufficient detail) and this should be a point of attention in the future.

CRedit authorship contribution statement

Christopher McCrum: Conceptualization, Methodology, Formal analysis, Investigation, Data Curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Jorg van Beek:** Investigation, Writing – review & editing. **Charlotte Schumacher:** Investigation, Writing – review & editing. **Sanne Jansen:** Investigation, Writing – review & editing. **Bas Van Hooren:** Investigation, Writing – review & editing. Authorship order after the lead author was determined by the time spent on article screening (most to least).

Declaration of Competing Interest

CM is a member of the ISPGR and a communications committee member of the ISPGR, for which Gait & Posture is the official society journal. CM and BVH have submitted to, published in and reviewed for Gait & Posture but are not involved with the editorial board or policies. The ISPGR and the editors and publisher of Gait & Posture had no input or influence on the design or reporting of this study.

Acknowledgements

The authors thank Cagla Fadillioglu for providing an older version of the author guidelines and Julie Stebbins (Deputy Editor of Gait & Posture) for providing information on the timing of the inclusion of the guideline.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.gaitpost.2021.12.010](https://doi.org/10.1016/j.gaitpost.2021.12.010).

References

- [1] D. Moher, S. Hopewell, K.F. Schulz, V. Montori, P.C. Gøtzsche, P.J. Devereaux, D. Elbourne, M. Egger, D.G. Altman, CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 340 (2010) c869, <https://doi.org/10.1136/bmj.c869>.
- [2] K. Dwan, T. Li, D.G. Altman, D. Elbourne, CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 366 (2019) 14378, <https://doi.org/10.1136/bmj.14378>.
- [3] M.K. Campbell, G. Piaggio, D.R. Elbourne, D.G. Altman, Consort 2010 statement: extension to cluster randomised trials. *BMJ* 345 (2012), e5661 <https://doi.org/10.1136/bmj.e5661>.
- [4] G. Piaggio, D.R. Elbourne, S.J. Pocock, S.J. Evans, D.G. Altman, Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *JAMA* 308 (24) (2012) 2594–2604, <https://doi.org/10.1001/jama.2012.87802>.
- [5] S.M. Eldridge, C.L. Chan, M.J. Campbell, C.M. Bond, S. Hopewell, L. Thabane, G. A. Lancaster, CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud.* 2 (2016) 64, <https://doi.org/10.1186/s40814-016-0105-8>.
- [6] J.P. Vandenbroucke, E. von Elm, D.G. Altman, P.C. Gøtzsche, C.D. Mulrow, S. J. Pocock, C. Poole, J.J. Schlesselman, M. Egger, Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLoS Med* 4 (10) (2007), e297 <https://doi.org/10.1371/journal.pmed.0040297>.
- [7] J.A. Betts, J.T. Gonzalez, L.M. Burke, G.L. Close, I. Garthe, L.J. James, A. E. Jeukendrup, J.P. Morton, D.C. Nieman, P. Peeling, S.M. Phillips, T. Stellingwerff, L.J.C. van Loon, C. Williams, K. Woolf, R. Maughan, G. Atkinson, PRESENT 2020: text expanding on the checklist for proper reporting of evidence in sport and exercise nutrition trials. *Int J. Sport Nutr. Exerc Metab.* 30 (1) (2020) 2–13, <https://doi.org/10.1123/ijsnem.2019-0326>.
- [8] S.D. Collins, N. Peek, R.D. Riley, G.P. Martin, Sample sizes of prediction model studies in prostate cancer were rarely justified and often insufficient. *J. Clin. Epidemiol.* 133 (2021) 53–60, <https://doi.org/10.1016/j.jclinepi.2020.12.011>.
- [9] L. Raittio, A. Launonen, V.M. Matilla, A. Reito, Estimates of the mean difference in orthopaedic randomized trials: obligatory yet obscure. *BMC Med Res Method.* 21 (1) (2021) 59, <https://doi.org/10.1186/s12874-021-01249-2>.
- [10] S. Tulka, B. Geis, C. Baulig, S. Knippschild, F. Krummenauer, Validity of sample sizes in publications of randomised controlled trials on the treatment of age-related macular degeneration: cross-sectional evaluation. *BMJ Open* 9 (10) (2019), e030312 <https://doi.org/10.1136/bmjopen-2019-030312>.
- [11] Guide for Authors. <https://www.elsevier.com/journals/gait-and-posture/0966-6362/guide-for-authors>. (Accessed 15 June 2021) 2021.
- [12] R.J. Prescott, Editorial: Avoid being tripped up by statistics: Statistical guidance for a successful research paper. *Gait Posture* 72 (2019) 240–249, <https://doi.org/10.1016/j.gaitpost.2018.06.172>.
- [13] T. Dreher, T. Wren, Introduction to special editorial on statistical guidance for a successful research paper. *Gait Posture* 72 (2019) 239, <https://doi.org/10.1016/j.gaitpost.2018.09.026>.
- [14] D. Lakens, Sample Size Justification. *PsyArXiv* (2021), <https://doi.org/10.31234/osf.io/9d3yf>.
- [15] C. McCrum, J. van Beek, B. Van Hooren, S. Janssen, C. Schumacher, Sample Size Justifications in Gait & Posture, Open Science Framework (2021), <https://doi.org/10.17605/OSF.IO/Z5C3R>.
- [16] C. McCrum, J. van Beek, B. Van Hooren, S. Janssen, C. Schumacher, Sample Size Justifications in Gait & Posture, Open Science Framework (2021), <https://doi.org/10.17605/OSF.IO/5V8WP>.
- [17] D. Martins Ribeiro, Comportamento do pé no período de balanço do ciclo da marcha de mulheres: estudo comparativo entre diferentes faixas etárias e sua relação com as quedas. Universidade de Brasília, Brasília (2018).
- [18] G.A.S. Bueno, Marcha de mulheres idosas e risco de quedas: influência do histórico de queda e medo de cair. Universidade de Brasília, Brasília (2019).
- [19] D. Knudson, Confidence crisis of results in biomechanics research. *Sports Biomech.* 16 (4) (2017) 425–433, <https://doi.org/10.1080/14763141.2016.1246603>.
- [20] A.R. Caldwell, A.D. Vigotsky, M.S. Tenan, R. Radel, D.T. Mellor, A. Kreutzler, I. M. Lahart, J.P. Mills, M.P. Boisgontier, Moving sport and exercise science forward: a call for the adoption of more transparent research practices. *Sports Med* 50 (3) (2020) 449–459, <https://doi.org/10.1007/s40279-019-01227-1>.
- [21] G. Abt, C. Boreham, G. Davison, R. Jackson, A. Nevill, E. Wallace, M. Williams, Power, precision, and sample size estimation in sport and exercise science research. *J. Sports Sci.* 38 (17) (2020) 1933–1935, <https://doi.org/10.1080/02640414.2020.1776002>.
- [22] T.W. Beck, The importance of a priori sample size estimation in strength and conditioning research. *J. Strength Cond. Res* 27 (8) (2013) 2323–2337, <https://doi.org/10.1519/JSC.0b013e318278eeaa>.
- [23] A.R. Caldwell, S.N. Cheuvront, Basic statistical considerations for physiology: The journal Temperature toolbox. *Temp. (Austin, Tex.)* 6 (3) (2019) 181–210, <https://doi.org/10.1080/23328940.2019.1624131>.
- [24] I. Halperin, A.D. Vigotsky, C. Foster, D.B. Pyne, Strengthening the Practice of Exercise and Sport-Science Research. *Int J. Sports Physiol. Perform.* 13 (2) (2018) 127–134, <https://doi.org/10.1123/ijspp.2017-0322>.
- [25] C. McCrum, Benefits of publishing power analysis reports. *J. Brief. Ideas.* (2021) <https://doi.org/10.5281/zenodo.4775072>.
- [26] C. Albers, D. Lakens, When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *J. Exp. Soc. Psychol.* 74 (2018) 187.
- [27] C.G. Moore, R.E. Carter, P.J. Nietert, P.W. Stewart, Recommendations for planning pilot studies in clinical and translational research. *Clin. Transl. Sci.* 4 (5) (2011) 332–337, <https://doi.org/10.1111/j.1752-8062.2011.00347.x>.
- [28] A.C. Leon, L.L. Davis, H.C. Kraemer, The role and interpretation of pilot studies in clinical research. *J. Psychiatr. Res* 45 (5) (2011) 626–629, <https://doi.org/10.1016/j.jpsychires.2010.10.008>.
- [29] D.G. Altman, J.M. Bland, Absence of evidence is not evidence of absence. *BMJ* 311 (7003) (1995) 485, <https://doi.org/10.1136/bmj.311.7003.485>.
- [30] M. Noordzij, G. Tripepi, F.W. Dekker, C. Zoccali, M.W. Tanck, K.J. Jager, Sample size calculations: basic principles and common pitfalls. *Nephrol. Dial. Transpl.* 25 (5) (2010) 1388–1393, <https://doi.org/10.1093/ndt/gfp732>.
- [31] A. McKeown, J.S. Gewandter, M.P. McDermott, J.R. Pawlowski, J.J. Poli, D. Rothstein, J.T. Farrar, I. Gilron, N.P. Katz, A.H. Lin, B.A. Rappaport, M. C. Rowbotham, D.C. Turk, R.H. Dworkin, S.M. Smith, Reporting of sample size calculations in analgesic clinical trials: ACTTION systematic review. *e191-197, J. Pain.* 16 (3) (2015) 199–206, <https://doi.org/10.1016/j.jpain.2014.11.010>.
- [32] B. Copesey, J.Y. Thompson, K. Vadher, U. Ali, S.J. Dutton, R. Fitzpatrick, S.E. Lamb, J.A. Cook, Sample size calculations are poorly conducted and reported in many randomized trials of hip and knee osteoarthritis: results of a systematic review. *J. Clin. Epidemiol.* 104 (2018) 52–61, <https://doi.org/10.1016/j.jclinepi.2018.08.013>.
- [33] T. Clark, U. Berger, U. Mansmann, Sample size determinations in original research protocols for randomised clinical trials submitted to UK research ethics committees: review. *BMJ* 346 (2013) f1135, <https://doi.org/10.1136/bmj.f1135>.
- [34] S. Tulka, S. Knippschild, S. Funck, I. Goetjes, Y. Uluk, C. Baulig, Reporting of statistical sample size calculations in publications of trials on age-related macular degeneration, glaucoma and cataract. *PLoS One* 16 (6) (2021), e0252640 <https://doi.org/10.1371/journal.pone.0252640>.
- [35] A. Reito, L. Raittio, O. Helminen, Revisiting the sample size and statistical power of randomized controlled trials in orthopaedics after 2 decades. *JBJS Rev.* 8 (2) (2020), e0079, <https://doi.org/10.2106/jbjs.Rvvr.19.00079>.
- [36] M.A. Robinson, J. Vanrenterghem, T.C. Pataky, Sample size estimation for biomechanical waveforms: Current practice, recommendations and a comparison to discrete power analysis. *J. Biomech.* 122 (2021), 110451 <https://doi.org/10.1016/j.jbiomech.2021.110451>.
- [37] P. Bacchetti, C.E. McCulloch, M.R. Segal, Simple, defensible sample sizes based on cost efficiency. *Biometrics* 64 (2) (2008) 577–585, <https://doi.org/10.1111/j.1541-0420.2008.01004.1.x>.
- [38] D.V. Knudson, Authorship and sampling practice in selected biomechanics and sports science journals. *Percept. Mot. Skills* 112 (3) (2011) 838–844, <https://doi.org/10.2466/17.Pms.112.3.838-844>.
- [39] J.S. Brach, S. Perera, S. Studenski, M. Katz, C. Hall, J. Verghese, Meaningful change in measures of gait variability in older adults. *Gait Posture* 31 (2) (2010) 175–179, <https://doi.org/10.1016/j.gaitpost.2009.10.002>.
- [40] M.J. Toebes, M.J. Hoozemans, S.E. Mathiassen, J. Dekker, J.H. van Dieen, Measurement strategy and statistical power in studies assessing gait stability and variability in older adults. *Aging Clin. Exp. Res* 28 (2) (2016) 257–265, <https://doi.org/10.1007/s40520-015-0390-8>.
- [41] D. Knudson, Twenty-year trends of authorship and sampling in applied biomechanics research. *Percept. Mot. Skills* 114 (1) (2012) 16–20, <https://doi.org/10.2466/11.Pms.114.1.16-20>.