

The Validity and Reliability of the My Jump Lab Artificial Intelligence Application

Lee A. Bridgeman^{1*}, Bailey Cameron¹ and James Steele¹

¹Department of Sport and Health, Solent University, Southampton, Hampshire, England.

*Corresponding Author: lee.bridgeman@solent.ac.uk

ABSTRACT

Jump height (JH) achieved in a countermovement jump (CMJ) has been suggested to allow for the monitoring of neuromuscular fatigue (NMF) and assessment of lower body power. Although force platforms (FP) are considered the gold standard for measuring CMJ height, they are expensive compared to mobile apps such as My Jump Lab (MJL). Therefore, this study aimed to assess the concurrent validity and agreement of the MJL app compared to a FP (ForceDecks [FD]) system and to determine its test-retest reliability. A convenience sample of 26 ($n = 11$ females and $n = 15$ males) recreationally active university sport students and staff (mean \pm SD; age: 23.08 ± 6.33 years; mass: 72.85 ± 9.93 kg; stature: 176.63 ± 10.18 cm) participated in the study. Participants attended the laboratory for testing on two separate occasions, separated by one week. After a standardised warm-up, they completed three CMJs on each occasion, with CMJ height simultaneously assessed by the FD and MJL app. The MJL Artificial Intelligence mode showed a mean bias of 4.32 cm [95% CI: 3.4, 5.26] overestimation with 95% limits of agreement ranging from -3.33 cm [95% CI: -4.96, -0.85] to 11.98 cm [95% CI: 10.13, 13.41]. Both methods demonstrated minimal mean bias (FD = 0.61 cm [95% CI: -0.31, 1.37] and MJL = 0.25 cm [95% CI = -0.48, 0.98]) between sessions, and both showed a similar width to their limits of agreement, ranging ~7 cm about the mean bias. In summary, the MJL overestimated CMJ height in this sample compared to the FD system, but both methods were reliable. Given the significant differences in cost for these two methods, teams on a budget may be interested in trialling the MJL app.

Keywords: Countermovement Jump, Force Decks, My Jump, Impulse-momentum

INTRODUCTION

A countermovement jump (CMJ) is performed by an individual starting in an upright standing position before initiating a downward squat movement followed immediately by an upward movement, which results in the athlete taking off (24). The jump height (JH) achieved in a CMJ has previously been suggested to allow for the monitoring of neuromuscular fatigue (NMF) (8, 9, 16) and assessment of lower body power (23, 26). As a result, this test is often utilised in laboratory and field-based testing. While several methods are available to assess CMJ height, force platforms (FP) are typically seen as the gold-standard measure (4) as they can capture both kinematic and kinetic variables. In particular, it has been proposed that dual FPs utilising the impulse-momentum (IM) relationship to assess CMJ height are desirable (14, 40). However, these systems come with a price tag (~£10,000+) that may be beyond the means of some teams and, therefore, is prohibitive. The proliferation of smartphone devices, which most coaches possess, now means they can access more affordable and easy-to-use solutions for assessing CMJ performance (4) through app-based solutions. One such app is My Jump Lab (MJL), available on Android and Apple devices via a subscription-based model (1 month = £4.99, 1 year = £34.99, Lifetime = £99.99 – prices correct at time of writing). A meta-analysis which investigated the validity of the app reported a correlation using a fixed effect model of 0.994 (95% CI = 0.992 – 0.995, $p < 0.001$) and 0.992 (95% CI = 0.998 – 0.995, $p < 0.001$) when using a random effect model to compare jump performance with the criterion method (FP) (13). When examining the app's reliability for the CMJ using a fixed effects model, an ICC of 0.969 (95% CI = 0.965 – 0.972, $p < 0.001$) was reported and 0.982 (95% CI = 0.961 – 0.992,

$p < 0.001$) when using a random effect model (13). The authors concluded that the MJL app was valid and reliable for assessing jump performance (13).

One issue with previous iterations of the MJL app was that it required practitioners to manually identify an athlete's take-off and touch-down from frame-by-frame video once the jumps had been completed. While this may not be an issue when working with individual athletes, it is suggested that when working with squads of 20-plus athletes, this may become a labour-intensive process for time-poor support staff (30). A newer version of the app has been launched to alleviate these issues. This app now allows real-time jump height (JH) measurement without post-processing (4). At present, two studies have investigated the reliability and validity of this app (4, 30). In the 1st proof of concept study (single participant), Balsalobre-Fernandez (4) completed 400 jumps over 24 consecutive weeks, recording JH simultaneously on a FP (Hawkins Dynamics) and the MJL app. They reported a very high correlation ($r = 0.971$, 95% CI = 0.963 – 0.975) and large agreement (ICC 0.969, 95 CI = 0.963 – 0.975) between the measures. Initially, though, large differences were reported between the instruments (mean absolute difference = 0.06 ± 0.01 m, $d = 4.4$, $p < 0.001$). However, after applying a regression equation to correct the app's raw data, non-significant, trivial differences were reported (mean absolute difference = 0.01 ± 0.008 m, $d = 0.1$, $p = 1.000$) between the devices. In a further study, Senturk et al. (30) compared the CMJ height recorded simultaneously by the MJL app and a FP (Vald – Force Decks) in 36 recreationally active participants. They reported a nearly perfect correlation ($r = 0.968$, $p = 0.001$) between the devices and very good agreement (mean difference = -1.016 cm, 95% CI = -1.229 – -0.803 cm) (30). This study also reported high intra-session reliability for the app (SEM = 0.43 cm; CV = 1.23%). These findings led the authors to conclude that the app was reliable and valid for measuring CMJ height (30).

However, it should be noted that in this study, all the participants were male and completed five jumps with a two-minute rest period between jumps (30). The authors suggest that this is different from what actually happens in a team sport environment, where there are often many athletes to test in a short period, and a 30-second rest period, as employed in this study, is more appropriate. Previous research has also noted concerns about replication in sport and exercise science; thus, further research using

similar procedures to see whether the previous results are replicated is considered valuable (25). Therefore, this study aimed to assess the concurrent validity and agreement of the MJL application for CMJ height compared to the ForceDecks system and its test-rest reliability.

METHOD

Experimental Approach to the Study

This cross-sectional study aimed to investigate the validity of the MJL application for measuring JH compared to the gold-standard method (jump height measured by the impulse-momentum calculation using a FP). The secondary aim was to investigate each piece of equipment's reliability. A convenient sample of recreationally active participants with previous CMJ experience was recruited for this study. Participants attended the laboratory for testing on two separate occasions, separated by one week, and completed three CMJs on each occasion, with JH simultaneously assessed by both the FP and MJL.

Participants

A convenience sample of 26 ($n = 11$ females and $n = 15$ males) recreationally active [As defined by McKay et al. (22)] university sports students and staff (mean ± SD; age: 23.08 ± 6.33 years; mass: 72.85 ± 9.93 kg; stature: 176.63 ± 10.18 cm) participated in the study. Before commencing testing, all participants were fully informed about the procedures, possible risks, and purpose of the study. All participants also completed a PAR-Q form and provided written informed consent. Participants were included if they met the criteria for being recreationally active (22) and excluded if they had suffered any lower-body injury in the 6 weeks before testing, had consumed caffeine two hours before testing or if their PAR-Q form identified any issues which may prevent them from performing maximal effort activities. The Solent University Ethics Committee approved this study.

Procedures

Participants were asked to refrain from lower body exercise 24 hours before each testing session. All testing took place at the same time of the day for all participants (± 1 hour) during both testing sessions. When the participants arrived at the laboratory, mass (Seca 875, Seca, Hamburg, Germany) and stature (Seca portable stadiometer, Seca, Hamburg,

Germany) were recorded. The participant's stature (cm) was recorded with shoes on to ensure measurement accuracy as they completed the jumps with shoes on as the app developer recommended. Before completing the jumps, each participant did a warm-up consisting of five minutes of cycling (Wattbike Pro, Wattbike, UK) at 100 W (80 – 90 rpm) followed by five bodyweight squats and three CMJs (30 seconds between jumps). Upon completion of the warm-up, the participants rested for three minutes before completing three CMJ trials (30 seconds between trials). All CMJ trials took place using the FDLite force plates (ForceDecks [FD], Vald, Brisbane, Australia) sampling at 1000 Hz. These FP have been reported to have high reliability (ICC = 0.93) when using the IM method to calculate CMJ height (24). A study by Collings et al. (10) also reported a 5% relative difference between FD and an embedded laboratory FP system and excellent test-retest reliability (ICC = 0.97 [0.92 - 0.99]) for JH. Before jumping, the FD were zeroed, and then the participants were weighed on the FD. After a period of quiet standing (~2 seconds), the participants squatted to a self-selected depth with their hands placed akimbo and were instructed to jump as high as possible for each of the three jumps. Jump height (cm) calculated by the IM relationship was the FD variable of interest during these trials. Simultaneously, a tripod with an iPhone 15 Pro Max (Apple, California, USA) was set up two meters away from the FD and at a height of one meters (to keep the participants in the bounding boxes created at a rate of 60 frames per second [FPS]) to record these jumps (consistent across all trials). These were recorded using the MJL app (v.4.2.8, 2024) with the artificial intelligence (AI) setting activated, which uses computer vision techniques to detect the participant's movement in each frame of the live video. (4) This allows JH to be measured in real-time. A qualified strength and conditioning coach (LB) with experience in CMJ testing administered the warm-up and all jumps.

Statistical Analysis

The present analysis was not pre-registered as we had no a priori hypotheses and, given the limited sample size due to resource constraints, was considered exploratory. However, since we have made our data open and focused on estimation, communicating these results is justified as they may contribute to future meta-analyses. Further, our sample size aligns with typical sample sizes in the field (17). Inferential statistics were treated as highly unstable local descriptions of the relations between

model assumptions and data to acknowledge the inherent uncertainty in drawing generalised inferences from single and small samples (3). For all analyses, we opted to take an estimation-based approach, which is typical when examining validity and reliability. That is, we provide point estimates and the uncertainty in them for the statistical parameters reported. Two sets of models were employed to explore the JHs recorded from CMJ trials: one to examine the agreement between the two methods and one to explore the test-retest reliability for each method. Given we had nested data whereby each participant provided three trials for each method on two separate testing sessions, we adopted a mixed effects limits of agreement approach (27). This allowed us to estimate mean bias for each method compared with the gold standard for agreement utilising both first and second session data and any test-retest bias for reliability, in addition to upper and lower limits of agreement providing 95% coverage probabilities for both agreement and test-retest reliability. Models were fit using the lme4 package and using Restricted Maximum Likelihood Estimation. For each of the models fit, we used nonparametric case-based bootstrapping resampling 10000 times at the individual participant level and refitting models to construct 95% quantile intervals for both the mean bias and limits of agreement estimates. Bootstrapping was performed using the lmeresampler package.

Agreement

For the estimation of bias and limits of agreement regarding the agreement between the gold-standard Force-Decks and impulse-momentum method with the MJL Artificial Intelligence mode, we utilised the following mixed effects model structure:

$$\begin{aligned} D_{ist} &= \mu + \alpha_i + \gamma_s + \epsilon_{ist} \\ \alpha_i &\sim N(0, \sigma_\alpha^2) \\ \gamma_s &\sim N(0, \sigma_\gamma^2) \\ \epsilon_{ist} &\sim N(0, \sigma_\epsilon^2) \end{aligned}$$

Where D_{ist} is the difference between measurements taken between the two methods (i.e., $y_{ist2} - y_{ist1}$), where the method indexed by 1 is the gold-standard and method indexed by 2 is the comparison method, for participant i during session s and for trial t . Here μ is the overall mean of the between method differences (i.e., the mean bias), α_i is the random effect for the i^{th} participant, γ_s is the random effect for the s^{th} session which is nested within participant, and ϵ_{ist} is the error term. The 95% limits of agreement can then be calculated as:

$$\mu \pm 1.96 \sqrt{\sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_\epsilon^2}$$

With the square root of the total variance providing an estimate of the standard deviation of the differences for use in the conventional Bland-Altman limits of agreement calculation. A model was fit for each method in comparison to the gold-standard.

Reliability

For estimation of bias and limits of agreement regarding the test-retest reliability between each test session for each method, we utilised the following mixed effects model structure:

$$\begin{aligned} D_{it}^* &= \mu^* + \alpha_i^* + \epsilon_{it}^* \\ \alpha_i^* &\sim N(0, \sigma_{\alpha^*}^2) \\ \epsilon_{it}^* &\sim N(0, \sigma_{\epsilon^*}^2) \end{aligned}$$

Where D_{it} is the difference between measurements taken between the two sessions for a given method (i.e., $y_{it2} - y_{it1}$), where the session indexed by 1 is the first test session and the session indexed by 2 is the second test session, for participant i and for trial t (note we use the superscript $*$ to distinguish this from the agreement model). Here μ^* is the overall mean of the between session differences (i.e., the mean bias), α_i^* is the random effect for the i^{th} participant, and ϵ_{it}^* is the error term. The 95% limits of agreement can then be calculated as:

$$\mu^* \pm 1.96 \sqrt{\sigma_{\alpha^*}^2 + \sigma_{\epsilon^*}^2}$$

with the square root of the total variance providing an estimate of the standard deviation of the differences for use in the conventional Bland-Altman limits of agreement calculation. A model was fit for each method to examine its test-retest reliability.

Open data, code, and materials

All data and code are presented in the supplementary materials (<https://osf.io/z9q2k/>). The *renv* package was used for package version reproducibility, and a function-based analysis pipeline using the *targets* package was employed (the analysis pipeline can be viewed by downloading the R Project and running the function `targets::tar_visnetwork()`). We cite all packages used in the analysis pipeline below using the *grateful* package (29):

We used R version 4.3.3 (32) and the following R packages: *glue* v. 1.7.0 (15), *janitor* v. 2.2.0 (11), *knitr* v. 1.47 (35-37), *lme4* v. 1.1.35.4 (6),

lmeresampler v. 0.2.4 (21), *patchwork* v. 1.2.0 (28), *quarto* v. 1.4 (1), *renv* v. 1.0.5 (33), *rmarkdown* v. 2.27 (2, 38, 39), *tarchetypes* v. 0.9.0 (18), *targets* v. 1.7.1 (19), *tidyverse* v. 2.0.0 (34).

RESULTS

Table 1 shows the mean \pm SD CMJ heights (cm) for each piece of equipment in both sessions.

Agreement

The mean bias and limits of agreement (LoA), along with their bootstrapped interval estimates, can be seen for the agreement between methods in Figure 1, along with the raw data. The MJL Artificial Intelligence mode showed a mean bias of 4.32 cm [95% CI: 3.4, 5.26] overestimation with 95% limits of agreement ranging from -3.33 cm [95% CI: -4.96, -0.85] to 11.98 cm [95% CI: 10.13, 13.41].

Reliability

The mean bias and LoA, along with their bootstrapped interval estimates, can be seen for the test-retest reliability of each method between sessions in Figure 2, along with the raw data. Both methods demonstrated minimal mean bias between sessions, each typically less than 1 cm, and both demonstrated a similar width to their LoA, ranging ~ 7 cm about the mean bias.

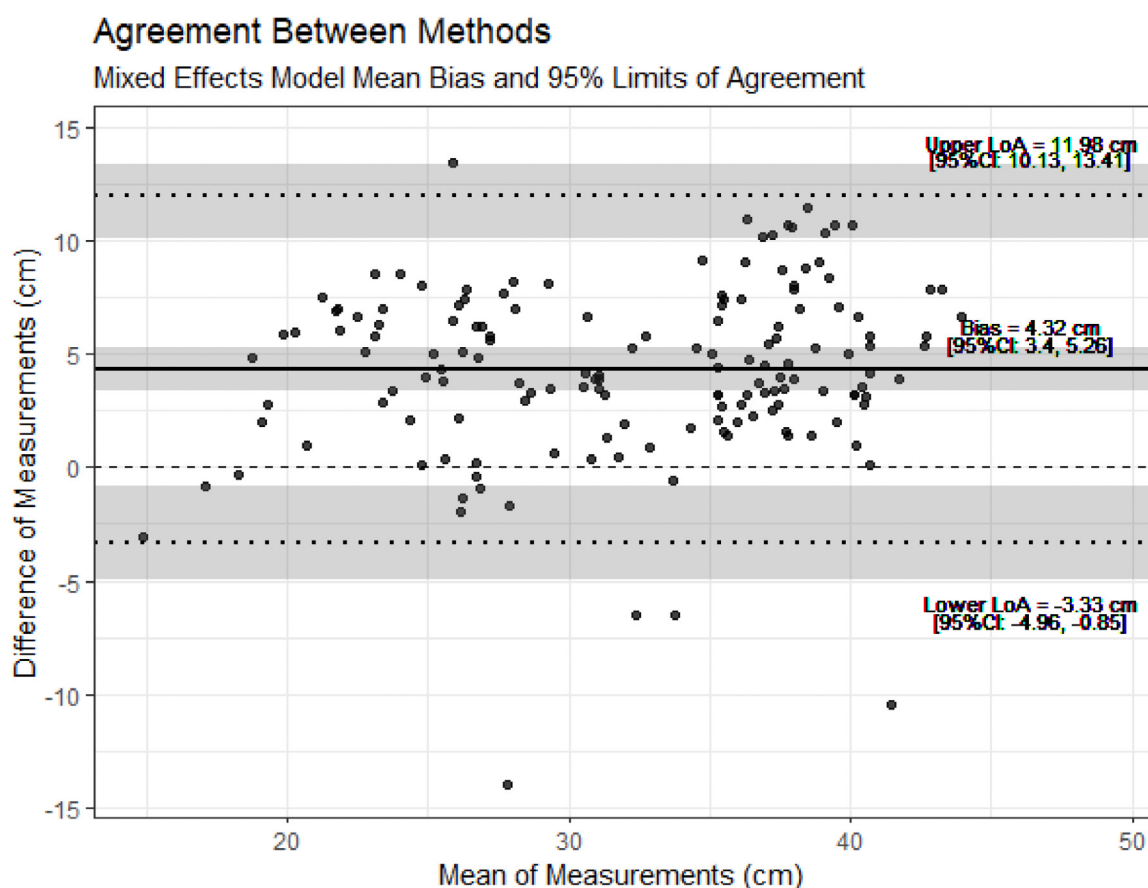
DISCUSSION

This study aimed to assess the concurrent validity and agreement of the MJL app for CMJ height compared to the FD system and evaluate its test-retest reliability. The main finding was that the MJL app overestimated CMJ height compared to the FD system (Mean bias = 4.32 cm [95% CI: 3.4, 5.26]). Concerning test-retest reliability, the bias for each device was found to be minimal (FD = 0.61 cm [95% CI: -0.31, 1.37] and MJL = 0.25 cm [95% CI: -0.48, 0.98]) and both showed a similar width to their LoA.

The current study agrees with Senturk et al. (30), who reported that the FD system produced lower JH than the MJL app (mean difference = -1.016 cm,

Table 1. Mean \pm SD jump height (cm) for FD and MJL of equipment in sessions 1 and 2.

Equipment	Session 1 Jump Height (mean \pm SD)	Session 2 Jump Height (mean \pm SD)
ForceDecks	29.68 \pm 7.03 cm	30.44 \pm 6.50 cm
My Jump Lab	34.13 \pm 7.63 cm	34.70 \pm 6.73 cm

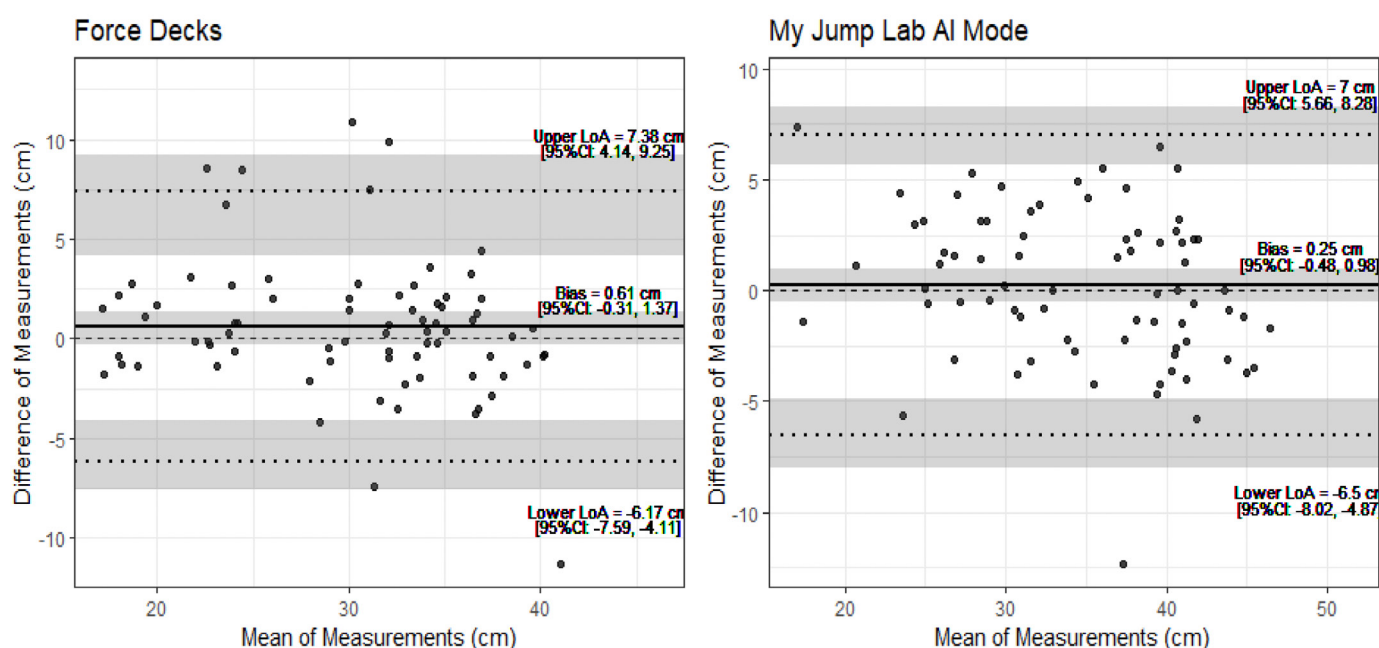


Intervals estimates for mean bias and limits of agreement are 95% quantiles from nonparametric bootstrap resampling participants 10000 times

Figure 1. Mean bias (thick horizontal line) and 95% limits of agreement (dotted horizontal lines) along with 95% quantile interval estimates for agreement of My Jump Lab Artificial Intelligence mode in comparison to the gold-standard (Force Decks impulse-momentum method).

Test-Retest Reliability Across Methods

Mixed Effects Model Mean Bias and 95% Limits of Agreement



Intervals estimates for mean bias and limits of agreement are 95% quantiles from nonparametric bootstrap resampling participants 10000 times

Figure 2. Mean bias (thick horizontal line) and 95% limits of agreement (dotted horizontal lines) along with 95% quantile interval estimates for each the test-retest reliability of method between sessions.

95% CI = -1.229 – -0.803 cm). However, it should be noted that the difference is larger in this study, with a mean bias of 4.32 cm in favour of the MJL app. Although the previous study (30) utilised the same equipment, direct comparisons are difficult as the authors did not report the method used to calculate JH from the FD system. Thus, whether they utilised the IM method or flight time to determine JH is unclear and may account for the difference. A further study by Tan et al. (31) reported that the MJL app with AI activated had a coefficient of variation (CV) of 4.12% and an ICC of 0.980 (95% CI = 0.972 – 0.986), suggesting the app is reliable in agreement with the current study. The agreement between a FP (Hawkins Dynamics) and the MJL app was reported to have an ICC of 0.973 (95% CI = 0.965 – 0.979), which the authors suggest is a strong concordance between the measurement methods (31). A Bland-Altman plot supported this, showing reasonable agreement with a minor average bias (31). Based on these findings and those of the current study, it is suggested that the MJL app with AI activated is reliable and has acceptable validity compared to FP systems. Previous versions of the MJL app, which used manual identification of take-off and touch-down to calculate JH, have been reported to be reliable (5, 7, 13). However, as stated previously, manually calculating JH for every jump can be time-consuming and does not allow for real-time feedback (4, 30). In agreement with previous research (4, 30, 31) investigating the MJL app using AI to detect JH, the current study found minimal bias between the two testing sessions. Therefore, it is suggested that the MJL app with AI activated can reliably measure JH.

While using a dual FP and the IM calculation to measure JH is considered the gold standard (14, 40), a considerable cost is associated with purchasing such a system (~£10,000+), which may make it unfeasible for many individuals and teams. Thus, an app-based system such as MJL may be attractive for its cost, ease of use, and portability. The ability to give athletes real-time feedback as they jump is also an attractive feature of the updated app. However, FPs can provide a wealth of kinetic and kinematic data (20) that is not possible using MJL, allowing practitioners to assess both the outcomes and the strategies used to achieve them. This may be useful when using the CMJ to assess NMF, as it has been suggested that athletes who are still not fully recovered can mitigate reductions in JH by altering their jumping strategy (e.g. longer eccentric duration) (12). Thus, if you only monitored JH, you may incorrectly decide that the athlete is

fully recovered, which could increase their injury risk. Therefore, when considering which system to purchase, ease of use, cost implications, and assessing what metrics you need to monitor CMJ performance is necessary.

LIMITATIONS

A key limitation of this study is the participant group, which included only sports students and staff (classified as recreational athletes). Consequently, the findings may be less applicable to elite athletes, who typically achieve higher jump performance. However, the results are likely relevant for recreational athletes on non-elite teams who may not have access to costly FP solutions, making the MJL app a potentially cost-effective alternative. Additionally, clinical practitioners who conduct CMJ performance assessments could benefit from the MJL app's affordability and ease of use, allowing clients to conveniently perform regular at-home tests and then share their progress with professionals for ongoing monitoring and assessment.

CONCLUSION

In conclusion, the MJL overestimated CMJ height in this study compared to the FD system. However, the app's test-retest reliability was found to be good. Given the relatively cheap cost of the app in comparison to FP systems, teams with smaller budgets may wish to investigate the use of the MJL app with AI activated to monitor performance enhancements following training interventions and to assess NMF if JH is the primary metric they are interested in. While the app's portability makes it attractive for coaches who want to evaluate their athletes in the field when using FPs may be problematic.

ACKNOWLEDGEMENTS

The authors wish to thank the participants for taking part in this study.

Author Contributions

LB, BC and JS conceived and designed the study. LB and BC acquired the data, interpreted the results, and drafted and revised the manuscript. JS completed the data analysis and contributed to interpreting the results and manuscript revisions. All

authors provided final approval of the version to be published.

CONFLICTS OF INTEREST

No other authors have a conflict of interest to declare.

FUNDING

This study received no specific funding in order to be completed.

ETHICAL APPROVAL

The Solent University Ethics Committee approved this study.

DATES OF REFERENCE

Submission - 15/07/2024

Acceptance - 12/11/2024

Publication - 11/07/2025

REFERENCES

- Allaire, J, Dervieux, C. quarto: R interface to "Quarto" markdown publishing system. 2024.
- Allaire, J, Dervieux, C, McPherson, J, et al. rmarkdown: Dynamic documents for R [Manual]. 2024.
- Amrhein, V, Trafimow, D, Greenland, S. Inferential statistics as descriptive statistics: There is no replication crisis if we don't expect replication. *The American Statistician* 73: 262-270, 2019.
- Balsalobre-Fernández, C. Real time estimation of vertical jump height with a markerless motion capture smartphone app: A proof-of-concept case study. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*: 17543371241227817, 2024.
- Balsalobre-Fernández, C, Glaister, M, Lockey, RA. The validity and reliability of an iPhone app for measuring vertical jump performance. *Journal of sports sciences* 33: 1574-1579, 2015.
- Bates, D, Macherl, M, Bolker, B, et al. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1-48, 2015.
- Bishop, C, Jarvis, P, Turner, A, et al. Validity and reliability of strategy metrics to assess countermovement jump performance using the newly developed smartphone application. *Journal of Human Kinetics* 83: 185-195, 2022.
- Bishop, C, Turner, A, Jordan, M, et al. A framework to guide practitioners for selecting metrics during the countermovement and drop jump tests. *Strength & Conditioning Journal* 44: 95-103, 2022.
- Claudino, JG, Cronin, J, Mezêncio, B, et al. The countermovement jump to monitor neuromuscular status: A meta-analysis. *Journal of Science and Medicine in Sport* 20: 397-402, 2017.
- Collings, TJ, Lima, YL, Dutailis, B, et al. Concurrent validity and test-retest reliability of VALD ForceDecks' strength, balance, and movement assessment tests. *Journal of Science and Medicine in Sport*, 2024.
- Firke, S. janitor: Simple tools for examining and cleaning dirty data. 2023.
- Gathercole, R, Sporer, B, Stellingwerff, T, et al. Alternative countermovement-jump analysis to quantify acute neuromuscular fatigue. *International journal of sports physiology and performance* 10: 84-92, 2015.
- Gençoğlu, C, Ulupinar, S, Özbay, S, et al. Validity and reliability of "My Jump app" to assess vertical jump performance: a meta-analytic review. *Scientific reports* 13: 20137, 2023.
- Heredia-Jimenez, J, Orantes-Gonzalez, E. Comparison of three different measurement systems to assess the vertical jump height. *Revista Brasileira de Medicina do Esporte* 26: 143-146, 2020.
- Hester, J, Bryan, J. glue: Interpreted string literals [Manual]. 2024.
- Hughes, S, Chapman, DW, Haff, GG, et al. The use of a functional test battery as a non-invasive method of fatigue assessment. *PloS one* 14: e0212870, 2019.
- Lakens, D. Sample size justification. *Collabra: psychology* 8: 33267, 2022.
- Landau, WM. tarchetypes: Archetypes for targets [Manual]. 2021.
- Landau, WM. The targets R package: A dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software* 6: 2959, 2021.
- Lombard, W, Reid, S, Pearson, K, et al. Reliability of metrics associated with a counter-movement jump performed on a force plate. *Measurement in Physical Education and Exercise Science* 21: 235-243, 2017.
- Loy, A, Steele, S, Korobova, J. Imeresampler: Bootstrap methods for nested linear mixed-effects models [Manual]. 2024.
- McKay, AK, Stellingwerff, T, Smith, ES, et al. Defining training and performance caliber: a participant classification framework. *International journal of sports physiology and performance* 17: 317-331, 2021.
- McMaster, DT, Gill, N, Cronin, J, et al. A brief review of strength and ballistic assessment methodologies in sport. *Sports Medicine* 44: 603-623, 2014.
- Merrigan, JJ, Strang, A, Eckerle, J, et al. Countermovement jump force-time curve analyses: reliability and comparability across force plate systems. *The Journal of Strength & Conditioning*

- Research 38: 30-37, 2024.
25. Mesquida, C, Murphy, J, Lakens, D, et al. Replication concerns in sports and exercise science: a narrative review of selected methodological issues in the field. *Royal Society Open Science* 9: 220946, 2022.
26. Nuzzo, JL, McBride, JM, Cormie, P, et al. Relationship between countermovement jump performance and multijoint isometric and dynamic tests of strength. *The Journal of Strength & Conditioning Research* 22: 699-707, 2008.
27. Parker, RA, Scott, C, Inácio, V, et al. Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. *BMC Medical Research Methodology* 20: 1-14, 2020.
28. Pedersen, TL. patchwork: The composer of plots [Manual]. 2024.
29. Rodriguez-Sanchez, F, Jackson, CP, Hutchins, SD, et al. Grateful: Facilitate Citation of R Packages. 2023.
30. Şentürk, D, Yüksel, O, Akyildiz, Z. The concurrent validity and reliability of the My Jump Lab smartphone app for the real-time measurement of vertical jump performance. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*: 17543371241246439, 2024.
31. Tan, EC, Onn, SW, Montalvo, S. Measuring Vertical Jump Height With Artificial Intelligence Through a Cell Phone: A Validity and Reliability Report. *The Journal of Strength & Conditioning Research* 38: e529-e533, 2024.
32. Team, RC. R: A language and environment for statistical computing [Manual]. R Foundation for Statistical Computing. 2024.
33. Ushey, K, Wickham, H. renv: Project environments [Manual]. 2024.
34. Wickham, H, Averick, M, Bryan, J, et al. Welcome to the tidyverse: *Journal of Open Source Software*, v. 4. doi 10: 1686, 2019.
35. Xie, Y. A comprehensive tool for reproducible research in R, in: *Implementing reproducible computational research*. V Stodden, F Leisch, RD Peng, eds.: Chapman and Hall/CRC., 2014.
36. Xie, Y. *Dynamic documents with R and knitr* Chapman and Hall/CRC, 2015.
37. Xie, Y. knitr: A general-purpose package for dynamic report generation in R [Manual]. 2024.
38. Xie, Y, Allaire, JJ, Golemund, G. *R markdown: The definitive guide*. Chapman and Hall/CRC., 2018.
39. Xie, Y, Dervieux, C, Riederer, E. *R markdown cookbook*. Chapman and Hall/CRC., 2020.
40. Xu, J, Turner, A, Comfort, P, et al. A systematic review of the different calculation methods for measuring jump height during the countermovement and drop jump tests. *Sports Medicine* 53: 1055-1072, 2023.