DOI: 10.1002/ksa.12515

LETTER TO THE EDITOR

Knee Surgery, Sports Traumatology, Arthroscopy WILEY

Letter to the Editor concerning 'The Calf Raise App shows good concurrent validity compared with a linear encoder in measuring total concentric work': Let's not compare apples to oranges

Dear Editor,

I read with interest the recently published article titled 'The Calf Raise App shows good concurrent validity compared with a linear encoder in measuring total concentric work' by Ashnai et al. [1] When well conducted, concurrent validation studies are important for clinicians and scientists to establish the relative interchangeability of devices and their outcome measures. While the title indicates overall favourable results, large systematic biases and relatively large differences between systems were noted (i.e., 25%). I am concerned that these differences are due to flaws in the data collection methods and thus I feel I need to caution readers regarding the quality of the data presented. Indeed, while the study aims to assess the concurrent validity of the Calf Raise App [5] compared to a linear encoder, I have identified several methodological issues and potential suboptimal use of the computer-vision-based application that may undermine the validity of the findings.

First, comparing the displacement of a linear position transducer placed on the heel to a marker placed on the lateral malleolus is fundamentally flawed to establish concurrent validity. The axis of rotation of the plantarflexion movement during heel raising is the metatarsal heads, and because the heel is further away from this rotation point than the malleolus, the former will always show greater linear displacement for the same angular displacement (Figure 1). This discrepancy is due to basic principles of angular kinematics, and thus, the two measurements cannot be considered equivalent for assessing concurrent validity. While the authors acknowledge this aspect in their discussion as 'the two instruments record heel raise height using different reference markers', the question arises as to why were the two then compared knowing that one would systematically be superior to the other? It would have made more sense to position both 'markers' at the same location, as well as to use a gold standard measurement for comparison of outcomes. According to Hurley et al. [8], concurrent validity refers to how well one measure is correlated with an existing gold standard measure, which is an important property to be established for new measures aiming to assess the same properties as an existing test. In the case of motion analysis, 3D motion is the established gold standard, not the linear encoder.

To further substantiate this methodological oversight, published [4] and unpublished data from our laboratory were pooled from studies where we have one 3D marker positioned below the lateral malleolus and one positioned on the heel. Positional data in 3D were recorded using retroreflective markers of 12.5 mm in diameter using an 8-camera Oqus 700 3D motion capture system sampling at 60 Hz and the Qualisys Track Manager (v.2019.1.4000; Qualisys AB) (Table 1). Data are from 23 males (mean ± standard deviation for age, height, mass, and body mass index: 32.7 ± 10.7 y, 178.7 ± 8.5 cm, 88.1 ± 21.2 kg and 27.7 ± 5.7 kg/m^2) and 19 females (29.3 ± 9.4 y, 165.5 ± 7.3 cm, 62 ± 8.3 kg and 22.8 ± 2.5 kg/m^2) each completing between one and six calf raise tests to fatigue with the foot starting from either the floor, on a 10° incline, or on the edge of a step [6]. From 210 distinct calf raise test performances, participants completed 30 ± 10 repetitions (range: 12–75). Data from Table 1 clearly shows a systematic bias wherein calf raise test outcomes are approximately 28% greater when a marker is placed on the heel versus below the lateral malleolus, but the two are almost perfectly correlated ($r \geq 0.925$). The approximate 25% difference reported by Ashnai et al. [1] between the linear encoder and the Calf Raise App is almost exclusively due to marker positioning. Of great concern, however, are their relatively poorer correlation coefficients (intra-class

Abbreviations: 2D, two dimensional; 3D, three dimensional.

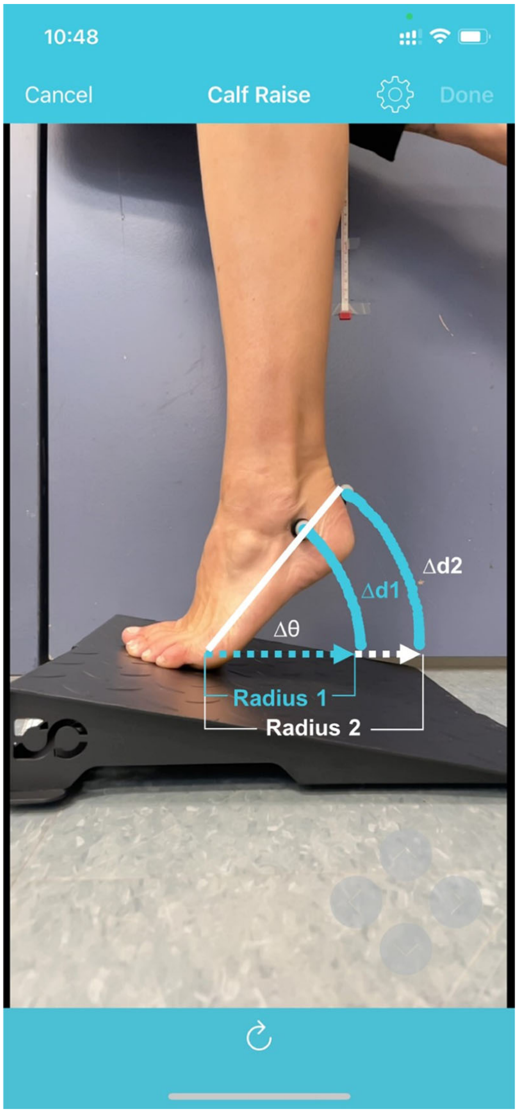

FIGURE 1 Calf Raise App screenshots of two markers: one positioned below the lateral malleolus and the other on the heel. The radius of each marker from the metatarsal heads is depicted. For the same angular displacement of the ankle joint (Δθ), the linear displacement is greater for the heel (Δd2) than for the malleolus (Δd1) marker due to the longer radius.

correlation 0.62–0.89) for average peak height and total positive displacement between systems, pointing to other sources of methodological errors described in the next paragraphs.

The second methodological concern relates to the fact the authors did not mention any calibration procedures for the Calf Raise App. The application relies on computer vision to track a circular marker of a known diameter [4]. The preset diameter in the application for calibration is of 2.4 cm, in line with the initial validation studies [4, 7]. There is a need for users of the Calf Raise App to adjust the preset diameter if a circular marker of a different diameter to 2.4 cm is used. Using appropriate calibration is essential for the valid interpretation of distances derived from 2D videos or else there will be systematic errors introduced [10]. However, the authors state, 'A small piece of black tape was placed on the distal part of the lateral malleolus', with no report on its diameter or length. Based on their image, the small piece of tape was not circular, appeared smaller than 2.4 cm, and was likely of an inconsistent diameter between participants. Based on a case study analysis of one participant completing 30 repetitions with a 2.4 cm circular marker placed below the lateral malleolus, the total displacement would be 365 cm when calibration is appropriate. If the scale is set to 3.2 cm (i.e., the marker is 25% smaller than the set calibration), the total displacement recorded is 280 cm (−33%). When reversed (i.e., the marker is 3.2 cm and 25% larger than the preset calibration of 2.4 cm), the total displacement reaches 494 cm (+35%). Therefore, not adjusting the calibration will impact Calf Raise App outcomes. Furthermore, the piece of tape was placed on the distal part of the lateral malleolus, which is usually curved. Even if a circular piece of tape was used, if it was flattened onto the curved aspect of the lateral malleolus, the circular shape would be distorted and the calibration would be negatively affected, which is why the marker should be placed on a flat surface below the malleolus. The use of a noncircular piece of tape (as appears to be the case for Ashnai et al. [1]) is problematic for the computer vision tracking algorithm. The app is designed to track a defined circular marker, and using an irregular shape could result in inconsistent tracking and further measurement errors (Figure 2). All these concerns and overt omission of reporting on calibration procedures suggest that the Calf Raise App was not properly scaled or calibrated, which is a critical step in 2D video motion analysis. The lack of valid calibration likely introduced significant measurement errors, further compromising the validity of findings. The fact that data from two participants registered a higher average than peak heel raise height and were disregarded points to the inappropriate calibration procedures, attention to best practices in videography [2], and overall use of the application.

Additionally, I am concerned about the potential inaccuracies in the linear encoder measurements and the distance of 1 m used to video record the calf raise motion. The authors did not state in their methods whether the linear encoder accounts for non-vertical displacements. If the encoder is not perfectly aligned to vertical, it would introduce errors that were not addressed in this study. It is not clear if the initial position of the transducer positioned on the heel of individuals defines zero or not, which will influence the interpretation of peak height outcomes. The authors' first mention of some of these aspects is in the discussion where they state that the linear encoder measures displacement in any direction from its attachment, not

TABLE 1 Outcomes from 210 calf raise tests performed by 42 noninjured participants.

Outcomes	Marker Heel	Malleolus	Difference (Heel–Malleolus) Raw units	Percentage (%)[a]	*p* Value[b]	Correlation Pearson *r* [95% CI]	*p* Value
Total positive displacement (cm)	354 ± 123	254 ± 91	99 ± 37	28.2 ± 4.4	<0.001	0.986 [0.982, 0.990]	<0.001
Total concentric work (J)	2545 ± 934	1824 ± 666	721 ± 297	28.2 ± 4.3	<0.001	0.987 [0.982, 0.990]	<0.001
Peak height (cm)[c]	13.0 ± 1.8	9.3 ± 1.5	3.7 ± 0.7	28.7 ± 4.6	<0.001	0.925 [0.903, 0.943]	<0.001
Average peak height (cm)[d]	11.8 ± 1.9	8.5 ± 1.6	3.3 ± 0.7	28.2 ± 4.4	<0.001	0.949 [0.933, 0.961]	<0.001

Note: Data derived from the vertical displacement of markers recorded using a 3D motion capture system.
[a]Percentage calculated as [(Heel–Malleolus)/Heel] × 100.
[b]Paired two-tailed *t*-test.
[C]Peak height reflects peak vertical displacement of any one repetition during the calf raise test, where the initial position of the marker is zero.
[d]Average peak height is the total positive displacement divided by the number of repetitions for comparability with Ashnai et al. [1].

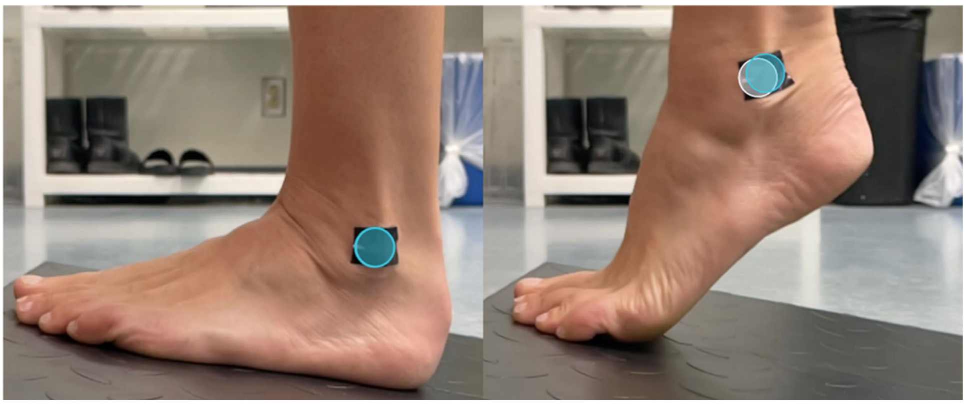

FIGURE 2 Illustrative example of the original circular calibration target applied in the application (left) being tracked at two separate locations on a rectangular piece of tape during the calf raise movement (right).

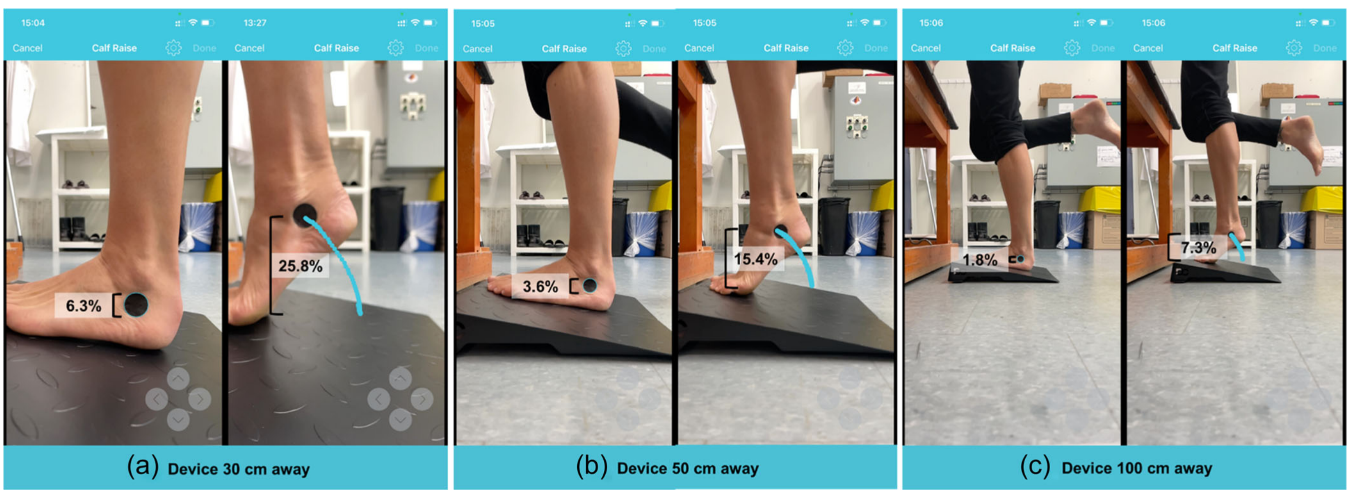

FIGURE 3 Influence of the recording device (iPhone 12 Pro) position on the relative representativeness of the 2.4 cm circular calibration target and of the calf raise motion in relation to the vertical video image size. Positions of the device are (a) 30 cm, (b) 50 cm, and (c) 100 cm from the foot during the calf raise test.

only vertical like the Calf Raise App. It appears that the linear encoder is the device introducing measurement error in the vertical plane. Moreover, the recording of the calf raise motion was conducted at a 1 m distance from the marker, which, combined with a small calibration target, would impair the precision and accuracy of the 2D video measurements. The use of larger calibration objects improves calibration precision and accuracy [3], and calibration objects should cover a considerable portion of the image [11]. If the 2D calibration target is too small, calibration accuracy declines [9]. Maximising the size of the marker and the motion of interest within the field of view increases the accuracy of digitisation [2], and this aspect needs consideration alongside maintaining the movement of interest central to the field of view to avoid lens distortion errors. Typically, distances of 30–50 cm achieve these goals for the Calf Raise App. As proof of concept, a recording of the calf raise motion was taken with a circular marker of 2.4 cm in diameter at three distances from the foot (Figure 3). The relative vertical size of the marker (that also acts as a calibration target) in relation to the corresponding vertical video image size decreases from 6.3% to 3.6% to 1.8% and the relative representation of the vertical calf raise movement decreases from 25.8% to 15.4% to 7.3% when the position of the recording device is moved from 30 to 50 to 100 cm (Figure 3). The accuracy of the marker sizing and resolution of the 100 cm set-up is relatively poor in comparison to closer set-ups and a source of considerable error.

The authors state that this study was designed as a validity study, yet it appears to be an add-on to another prospectively registered study (NCT05323773) with no initial intention of concurrent validation of the Calf Raise App. To properly assess the validity of the systems, similar reference points (e.g., the lateral side of the foot) should be used, and outcomes should be compared to a gold standard, such as 3D motion analysis.

To conclude, the methodological flaws in data collection and analysis significantly limit the validity of the inferences drawn from this study. I urge caution to readers in interpreting these findings from Ashnai et al. [1] given the probable suboptimal use of the Calf Raise App, especially in relation to Calf Raise App calibration and position of the recording device. Future well-conducted concurrent validation studies with proper use of the Calf Raise App and considerations of proper 2D video analysis methodologies are encouraged to address the current issues identified and ensure more accurate and reliable results. I encourage readers and potential Calf Raise App users to consult the series of five instructional videos created to maximise data quality: https://youtube.com/@calfraiseapp?si=XYAVvKJKHBuITz6Q.

I thank you for considering my comments.

Respectfully,

Kim Hébert-Losier

CONFLICT OF INTEREST STATEMENT

Kim Hébert-Losier is one of the developers of the free-to-use Calf Raise application and receives no financial incentive from downloads or use of the Calf Raise application.

ETHICS STATEMENT

The participant data included in this paper were collected as part of a project granted ethical approval from the health research ethics committee at the University of Waikato (HREC2020#11).

Kim Hébert-Losier

Division of Health, Engineering, Computing and Science, Te Huataki Waiora School of Health, University of Waikato, Tauranga, New Zealand

Correspondence

Kim Hébert-Losier, Division of Health, Engineering, Computing and Science, Te Huataki Waiora School of Health, University of Waikato, Tauranga, New Zealand.
Email: kim.hebert-losier@waikato.ac.nz

ORCID

Kim Hébert-Losier http://orcid.org/0000-0003-1087-4986

TWITTER

Kim Hébert-Losier @KimHebertLosier

REFERENCES

1. Ashnai, F., Lindskog, J., Brorsson, A., Nilsson-Helander, K. & Beischer, S. (2024) The Calf Raise App shows good concurrent validity compared with a linear encoder in measuring total concentric work. *Knee Surgery, Sports Traumatology, Arthroscopy*, 32, 2170–2177. Available from: https://doi.org/10.1002/ksa.12190
2. Bartlett, R. (2014) *Introduction to Sports Biomechanics: Analysing Human Movement Patterns*, 3rd edition. New York, NY: Routledge.
3. Boutros, N., Shortis, M.R. & Harvey, E.S. (2015) A comparison of calibration methods and system configurations of underwater stereo-video systems for applications in marine ecology. *Limnology and Oceanography: Methods*, 13, 224–236. Available from: https://doi.org/10.1002/lom3.10020
4. Fernandez, M.R., Athens, J., Balsalobre-Fernandez, C., Kubo, M. & Hébert-Losier, K. (2023) Concurrent validity and reliability of a mobile iOS application used to assess calf raise test kinematics. *Musculoskeletal Science & Practice*, 63, 102711. Available from: https://doi.org/10.1016/j.msksp.2022.102711
5. Hébert-Losier, K. & Balsalobre-Fernandez, C. (2020) Calf Raise [Mobile application software]. *Open Science Framework*. Available from: https://doi.org/10.17605/OSF.IO/N9SG7
6. Hébert-Losier, K., Fernandez, M.R., Athens, J., Kubo, M. & O'Neill, S. (2024) A randomised crossover trial on the effects of foot starting position on calf raise test outcomes: Position does matter. *The Foot*, 60, 102112. Available from: https://doi.org/10.1016/j.foot.2024.102112
7. Hébert-Losier, K., Ngawhika, T.M., Gill, N. & Balsalobre-Fernandez, C. (2022) Validity, reliability, and normative data on calf muscle function in rugby union players from the Calf Raise application. *Sports Biomechanics*, 1–22. In press. Available from: https://doi.org/10.1080/14763141.2022.2118158
8. Hurley, W.L., Denegar, C.R. & Hertel, J. (2011) *Research Methods: A Framework for Evidence-Based Clinical Practice*. Baltimore, MD: Lippincott Williams & Wilkins.
9. Liu, Z., Li, F., Li, X. & Zhang, G. (2015) A novel and accurate calibration method for cameras with large field of view using

combined small targets. *Measurement*, 64, 1–16. Available from: https://doi.org/10.1016/j.measurement.2014.11.027
10. Martin, T., Frisch, K. & Zwart, J. (2020) Systematic errors in video analysis. *The Physics Teacher*, 58, 195–197. Available from: https://doi.org/10.1119/1.5145415
11. Samper, D., Santolaria, J., Brosed, F.J., Majarena, A.C. & Aguilar, J.J. (2013) Analysis of Tsai calibration method using two- and three-dimensional calibration objects. *Machine Vision and Applications*, 24, 117–131. Available from: https://doi.org/10.1007/s00138-011-0398-9