Concurrent Validity Analysis of 2D Convolutional Mesh Regression Software Against Gold Measure While Measuring Calf Girth

B. Zepeda-Almazan[1] [1]Sheffield Hallam University | MSc Sports Engineering

Submitted: April. 22th, 2020.

Introduction

Recent studies have applied computer modelling in calculating key anthropometric variables that correlate to success in sport. Applications include using wingspan, arm girth, and calf girth to predict swimming times, throwing velocity, and rowing success respectively. Such tools allow strategists to enhance recruiting programs, training regimens, and dissect performance analysis. Mikulic et. al (2008) showed anthropometric inclusive models best predict elite rowers' performance ($R^2 = 0.85$) compared to solely physiological ($R^2 = 0.80$). Predictive software, Graph Convolutional Mesh Regression (Graph-CMR) uses a learning based monocular approach to report anthropometric variables through parameter regression from 2D imaging. Kolotouros et. al (2019) showed the program to have improved the reconstructed model error of Lassner et. al (2017) by 46.6%. Although progressive, the system's validity is challenged due to simplification of mesh vertices through non-parametric model fitting. The system's reliability is critically considered across various body types to mitigate the impact of premature application in elite and various sports.

Purpose

This study will use a gold measure to analyse the validity of Graph-CMR in measuring calf girth. The findings will provide critical correlation analysis and a priori criteria of $\alpha < 0.05$.

Procedure

Six participants were recruited (mean ± standard deviation: age: 22.3 ± 0.81 years; stature: 1.71 ± 0.84 m; mass: 69.18 ± 7.50 kg) and gave consent to be included in the study. Ethical consideration followed Sheffield Hallam guidelines. Participants were asked to wear tight fitting clothing exposing the lower leg and no shoes. Anonymity was ensured through participant male/ female numbered identifiers. Using a commercially available phone (One Plus 6t - 20MP), the participants' photos were taken against a bare wall in a well-lit space, (Figure 1) and uploaded to the Graph-CMR server. The participants calves were also measured using gold standard procedures dictated by a certificated anthropometrist. Calf girths were measured using a tape measure and were averaged over three repeated trials. Noted limitations are low sample size and potentially disruptive backgrounds.

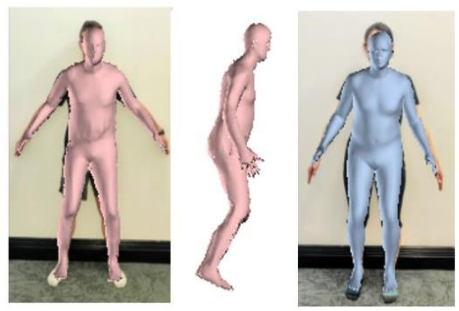

Figure 1
Visual output of Male & Female Participants through Graph-CMR software. Clear limitations in accounting for gender anatomical points are evident via fit of model.

Results

The results in Table 1 are expressed as means with statistical significance considered for values $p < 0.05$. Pearson's r coefficient was calculated ($r = -0.13$, $p = 0.81$), (ICC = -0.07, RMSE = 0.025). With an average model error of 6.1% (Female = 9.8%, Male = 2.8%). Figure 2A-1 shows a Bland Altman Plot (Bland & Altman, 1999) between the two measures with limits of agreement at the 95% confidence interval [0.02, -0.06], with clear proportional bias which was quantified by grouping the datum via gender and normalising against BMI via linear regression. The Bland Altman plot was iterated with corrected data values (Figure 2A-2) and summarised in Figure 2B, limits of agreement improved

Table 1 Compiled Raw Data | Measured in [m]. Proportional bias was normalised and included in data analysis as Normalised Model values.

ID	BMI	Gold Measure	Graph CMR	Normalised Model	Model Error [%]	Bias Corrected Error [%}
F1	23.29	0.35	0.39	0.35	+ 12.17	- 0.77
F2	23.33	0.36	0.39	0.35	+ 8.62	- 1.85
F3	23.40	0.37	0.40	0.37	+ 6.95	- 0.06
M1	23.50	0.40	0.39	0.40	- 1.21	+ 1.32
M2	24.03	0.40	0.39	0.39	- 2.09	- 2.75
M3	24.54	0.38	0.40	0.38	+ 5.18	+ 1.20

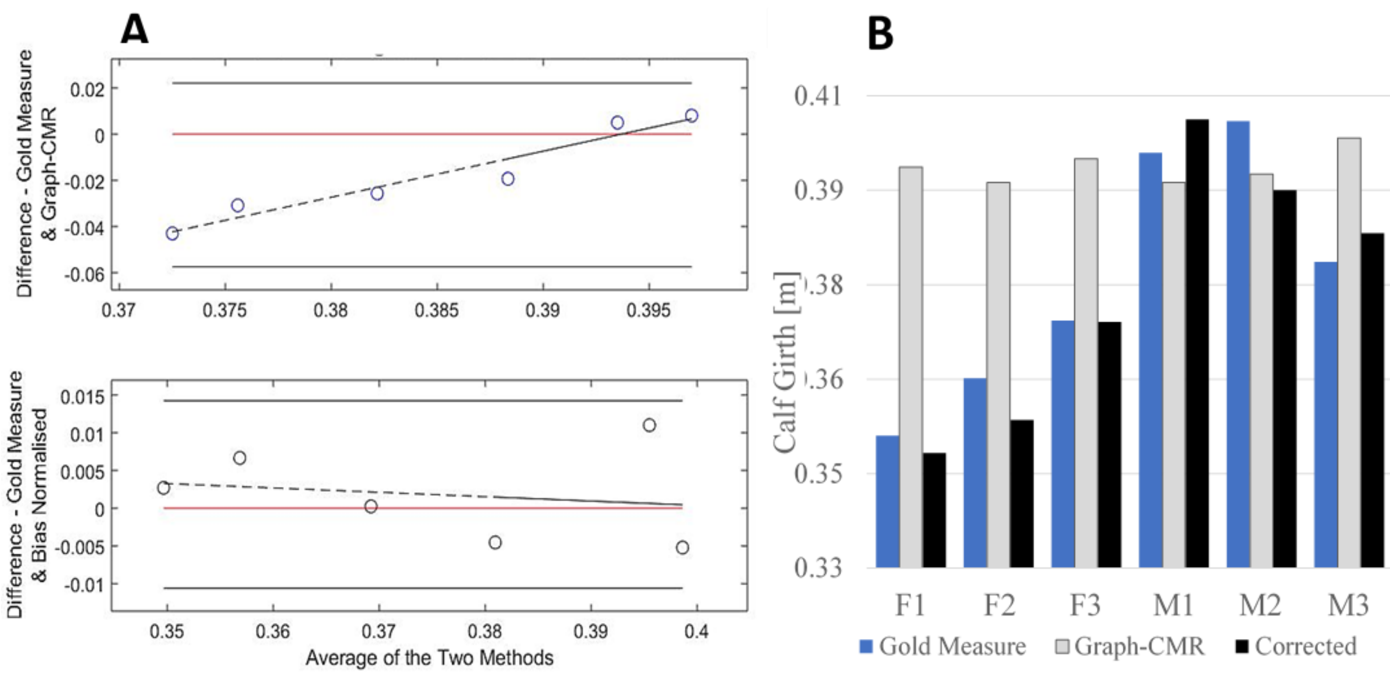

Figure 2 [A] Bland Altman plot of raw and corrected data; further analysis of the Bland Altman plot led to the normalisation of raw datum and found the best linear adjustment to be in accounting for BMI and gender, as reported in the findings. The limits of agreement raw data (95% CI): [0.02, -0.06] were improved to [0.015, -0.010] showing greater agreement and lowered levels of variance. [B] Collected data across measurement techniques | Qualitative analysis of datum highlights proportional bias and impacts of normalisation against BMI when grouped by gender.

to [0.015, - 0.010]. Pearson's r for corrected datum is ($r = 0.95$, $p < 0.01$), (ICC = 0.95, RMSE = 0.006), with an average model error 1.3% (Female = 0.9%, Male = 1.8%).

Conclusion

The goal of this study was to investigate the validity of Graph-CMR against a gold measure in measuring calf girth. Initial results show low accuracy with an evident proportional bias. The corrected datum shows strong agreement to the gold measure. Graph-CMR shows potential in rapid calculations and accessibility but requires rectification to high levels of error, potentially accountable by gender and BMI, before deployment to coaches and strategists. Further research should investigate quantifying bias and improving mesh refinement for unique anatomic markers.

Word Count: 550

References

Bland, J., & Altman, D. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160. doi: 10.1191/096228099673819272

Kolotouros, N., Pavlakos, G., & Daniilidis, K. (2019). Convolutional Mesh Regression for Single-Image Human Shape Reconstruction. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. doi: 10.1109/cvpr.2019.00463

Lassner, C., Romero, J., Kiefel, M., Bogo, F., Black, M. J., & Gehler, P. V. (2017). Unite the People: Closing the Loop Between 3D and 2D Human Representations. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi: 10.1109/cvpr.2017.500

Mikulić, P., & Ružić, L. (2008). Predicting the 1000m rowing ergometer performance in 12–13-year-old rowers: The basis for selection process? *Journal of Science and Medicine in Sport*, *11*(2), 218–226. doi: 10.1016/j.jsams.2007.01.008