

1           **Assessing quadriceps strength in patellofemoral pain**  
2           **patients: A study on the reliability and validity of a low-**  
3           **cost load-cell for clinical practice**

4   Germari Deysel, Mariaan van Aswegen, Mark Kramer\*

5   Physical Activity, Sport, and Recreation (PhasRec) Research Focus Area, North-West

6   University, Potchefstroom, North West Province, South Africa

7

8   \*Corresponding author

9   E-mail: [mark.kramer@nwu.ac.za](mailto:mark.kramer@nwu.ac.za) (MK)

10

11

## 12 **Abstract**

## 13 **Background**

14 Patellofemoral pain (PFP) is a common knee complaint affecting diverse populations both  
15 acutely and chronically. Quadriceps muscle weakness is one possible aetiology, but current  
16 devices for measuring muscle strength (isokinetic dynamometer and hand-held  
17 dynamometers) are frequently too expensive for practitioners, especially in under-resourced  
18 settings. There is a need to evaluate a low-cost device to manage rehabilitation of people with  
19 PFP.

## 20 **Methods**

21 Isometric quadriceps strength of participants aged 18-35 years (total [n = 33], control group  
22 [n = 17] and PFP group [n = 16]) were evaluated on an isokinetic dynamometer and a low-  
23 cost load cell at baseline and after an 8-week non-standardised intervention for validity  
24 scores.

## 25 **Results**

26 The load cell showed high absolute and relative reliability (intraclass correlation coefficient =  
27 0.89-0.99; typical error of measurement = 3.9-10.4%). Clinically meaningful difference  
28 scores (12.2-45 Nm) were greater than the typical error of measurement, implying sufficient  
29 sensitivity of the load cell to measure true changes in isometric quadricep strength. Strong to  
30 very strong correlations were evident between the load cell and isokinetic dynamometer  
31 torque measurements ( $r = 0.88-0.90$ ,  $SEE = 0.05-0.07$  Nm), but slope values ( $\beta = 0.65-0.77$ )  
32 indicated that torque from the load cell was typically lower than that obtained from the  
33 isokinetic dynamometer. An average systematic bias of 16.3-28.8 Nm was evident in favour

34 of the isokinetic dynamometer, with no statistically significant between-group differences  
35 being noted between the baseline and follow-up testing.

## 36 **Conclusion**

37 The load cell is a reliable instrument, sensitive enough to detect clinically meaningful  
38 differences in quadriceps strength in healthy individuals and those with PFP. The load cell  
39 lacks validity and cannot replace isokinetic dynamometry. Given the low cost and excellent  
40 reliability, the load cell can be a valuable tool to assess quadriceps muscle deficits and track  
41 rehabilitation progress in people with PFP.

42

## 43 **Introduction**

44 Patellofemoral pain (PFP) is a common knee complaint characterised by retropatellar pain  
45 (pain behind the kneecap) or peripatellar pain (pain around the kneecap), aggravated by at  
46 least one knee loading activity during weight bearing on a flexed knee (e.g., squatting, stair  
47 ambulation, jogging, running, hopping or jumping) [1]. Prevalence rates of PFP vary  
48 according to age, gender, and activity levels with the literature showing rates of 28% in  
49 adolescents [2], 23% in adults [3], 15% in female adults [4] and 40% in recreational runners  
50 [5]. People living with PFP often struggle with acute and chronic effects such as physical,  
51 emotional and social problems during sports participation, activities related to work and  
52 activities of daily living, which can persist several years [6–10]. Furthermore, PFP is  
53 considered a risk factor for the development of patellofemoral osteoarthritis [11], which has  
54 been attributed to reduced quadriceps strength relative to task-related loading of the  
55 patellofemoral joint [5,12,13]. It is therefore plausible to consider quadriceps strength as a  
56 *protective agent* against patellofemoral osteoarthritis cartilage loss [14].

57 Given that people with PFP can be divided into different subgroups, depending on the  
58 individual aetiologies [15], it is important to determine whether quadriceps weakness is part  
59 of the cause for PFP in the individual, and to track whether progress is made when  
60 conducting an exercise program to improve quadriceps strength [15]. Presently, the ‘gold  
61 standard’ instrument for the evaluation of quadriceps strength is the isokinetic dynamometer  
62 (ID) [16,17], which is beyond the affordability of most clinicians (~R500k-R1.50mil).  
63 Although a more reasonable alternative to the ID exists, such as a handheld dynamometer  
64 (HHD) [17], the reliability of an HHD depends on the strength and aptitude of the practitioner  
65 [18,19]. The latter is especially true when assessing stronger individuals or muscle strength  
66 around larger joints such as the knee joint [17] which may have important clinical  
67 implications for both patients and practitioners. Furthermore, HHDs are still relatively  
68 expensive (~R20k) when considering starting a rehabilitation practice, especially in under-  
69 resourced communities and clinical settings. Therefore, the purpose of this study was to: (i)  
70 establish the validity and reliability of a low-cost load cell (LC) (~R2k) for the assessment of  
71 isometric quadriceps muscle strength in people with PFP, (ii) evaluate the clinically  
72 meaningful difference (CMD) needed to enhance practitioner-based decision-making, and  
73 (iii) assess differences in isometric quadriceps strength with the use of a low-cost device  
74 between those with PFP and healthy control following an 8-week intervention.

75

## 76 **Materials and Methods**

### 77 **Study design**

78 This study used a repeated measures mixed study design consisting of a combination of  
79 between and within subject factors [20]. The between-subjects independent variable was  
80 group allocation (i.e., control group and PFP group), where group assignment was not

81 randomised (due to presence/absence of PFP). Within-subjects independent variables  
82 included time (repeated measures) and extremity (injured [or non-dominant] and uninjured  
83 [or dominant] limb).

## 84 **Participants**

85 A total of 35 participants volunteered for the study, of which 17 were part of the control  
86 group (female [n=10] and male [n=7]) and 18 were part of the PFP group (female [n=15] and  
87 male [n=3]). Given the requirements of the study design, a minimum total sample size of 24  
88 was calculated based on a repeated measures analysis of variance (ANOVA) design that  
89 incorporated (i) a within-between interaction, (ii) a moderate effect size ( $f = 0.25$ ), (iii) a  
90 type-I error rate of 5% ( $\alpha = 0.05$ ), (iv) a type-II error rate of 20% ( $\beta = 0.20$ ), (v) 2 groups  
91 (control [n = 17] and PFP [n = 17]), (vi) 2 repeated measurements, (vii) an anticipated  
92 dropout of 10%, and (viii) a minimum expected correlation of 0.50 among repeated  
93 measurements [21]. Accounting for a potential drop-out of 20%, the minimum sample size  
94 for adequate statistical power was 29 participants. Two participants were lost to follow-up,  
95 both in the PFP group (female [n=1] and male [n=1]), resulting in a final sample size of 33  
96 participants.

97 Participants were recruited by dispersing electronic flyers via social media platforms (e.g.,  
98 Facebook, WhatsApp), inviting prospective participants to contact the researcher for an  
99 information letter which outlined the details of the study. Participants who were willing to  
100 take part in the study were screened for eligibility (see inclusion and exclusion criteria below)  
101 and received an informed consent form to read through and were given 72 hours to sign and  
102 submit the form to an independent person affiliated with the study.

103

104 The inclusion criteria for the PFP group of this consisted of the following: (i) aged between  
105 18 and 35 years, (ii) could be male or female, (iii) had to have retropatellar and/or peripatellar  
106 pain aggravated by at least one activity that loads the patellofemoral joint during weight  
107 bearing on a flexed knee, such as squatting, stair climbing, jogging/running, and  
108 hopping/jumping, and (iv) had to participate in some form of rehabilitation program. The  
109 control group had similar inclusion criteria, with the exception for points (iii) and (iv).

110 The specific exclusion criteria for both the PFP and control group consisted of the following:  
111 participants should not (i) have had previous patellar dislocation or subluxation, (ii) have had  
112 previous injury or surgery to the knee, and (iii) have had recent (within the last 6 months)  
113 injury to the lower limbs (ex. Achilles tendinopathies, ankle sprain, etc.). All data collection  
114 occurred between 9<sup>th</sup> January 2023 to 22<sup>nd</sup> September 2023.

115 Ethics approval was granted by the Human Research Ethics Committee of the university  
116 (NWU-00163-22-A1), and all participants completed the informed consent forms prior to  
117 participation. An independent researcher not affiliated with the study served as an  
118 independent witness and collected the signed consent forms. All ethical procedures  
119 conformed to the requirements of ethical conduct set forth in the Declaration of Helsinki.

## 120 **Instruments**

121 A general demographic questionnaire was used to obtain the participant's contact details, e-  
122 mail address, age, sex, involved limb, dominant limb, previous injuries and PFP symptoms.  
123 The anterior knee pain scale (AKPS), also known as the Kujula patellofemoral scoring  
124 system, was completed electronically to capture knee-related pathologies where the total  
125 score out of 100 was captured. The AKPS was the chosen questionnaire, as it was developed  
126 specifically for evaluation of pain and disability in individuals with PFP [22], showed high  
127 test-retest reliability (intraclass correlation coefficient [ICC] = 0.95), and exhibited moderate

128 responsiveness to clinical change, which implies that the score will likely reflect meaningful  
129 changes in a patient's condition over time [23]. Body mass and stature were measured to  
130 calculate body mass index (BMI) [24]. Body mass was measured with an electronic scale  
131 (Seca 874, Seca, Germany) to the nearest 0.1 kg and stature with a portable stadiometer  
132 (Holtain Ltd., U.K.) to the nearest 0.01 m. A cycle ergometer (Wattbike Pro, Wattbike Ltd,  
133 Nottingham, UK) was used for a ten-minute warm-up before the testing on the ID for optimal  
134 muscle performance and reduced risk of injury. The ID (Cybex II, CSMi, Stoughton, MA,  
135 USA) was used as the gold standard for evaluating isometric muscle strength and a LC  
136 (Crane & Hanging Scale, Micro Mini CS300, Border Scales & Labels) served as the low-cost  
137 alternative to measure isometric quadriceps strength. Strength was evaluated with the knee  
138 joint at 60° of knee flexion, where peak force of the quadriceps is usually generated [25]. A  
139 digital goniometer (EasyAngle, Meloq AB, Stockholm, Sweden) was used to measure the  
140 knee angle in all instances to ensure true validity and replicability.

## 141 **Procedures**

142 The demographic questionnaire was completed first, and only at the baseline testing.  
143 Thereafter the participants completed the AKPS questionnaire verbally, where the score was  
144 calculated out of 100 and recorded electronically. Stature and body mass were measured  
145 during barefoot standing and with minimal clothing. Lower leg length was measured from the  
146 lateral condyle of the femur to the lateral malleolus of the tibia to use as the lever length in  
147 the formula to convert force to torque for the LC measurements. Prior to testing participants  
148 completed a ten-minute warm-up on a cycle ergometer at lowest resistance and a comfortable  
149 speed (rating of perceived exertion [RPE] < 2 on the modified Borg scale). Participants  
150 completed baseline testing in two sessions separated by 24 hours, and then repeated the same  
151 testing eight weeks later. During the first session, isometric strength of the quadriceps  
152 muscles was measured first with the ID, and thereafter with the LC. During the second

153 session, only the testing on the LC was completed for the reliability analysis. Reliability  
154 analysis was done both at baseline testing and at follow-up testing conducted after 8 weeks to  
155 ensure a stronger reliability score.

156 For the evaluation of isometric quadriceps strength on the ID, all variables related to the set-  
157 up were recorded to replicate the exact position for follow-up testing. The dominant limb  
158 (defined as the preferred kicking limb) in the control group and the uninjured limb in the PFP  
159 group were measured first. The participants were placed in a seated position, stabilized with  
160 upper body straps and an upper leg strap just above the knee joint. The knee was positioned at  
161 a 60° angle, using the machine angle provided by the dynamometer. The lateral femoral  
162 epicondyle was aligned with the axis of rotation of the dynamometer, and the resistance pad  
163 was positioned anterior to the distal tibia just superior to the lateral and medial malleoli. A  
164 gravity correction was performed to account for any potential additional torque induced on  
165 the attachment. The participant completed a warm-up round consisting of three repetitions,  
166 with the instruction to do one repetition at approximately 25%, 50% and 75% of perceived  
167 maximum effort before commencing of the test to assist with familiarisation of the test. A 10  
168 second rest period was allotted between trial repetitions. A one-minute rest period was  
169 granted after the warm-up, after which the participant completed repetitions of five second  
170 maximal extension contraction with ten seconds of rest between repetitions. Verbal  
171 encouragement was given throughout the whole procedure. Pain levels were monitored by  
172 asking the participant to rate pain on a scale of 1-10 throughout the procedure (where 1=pain  
173 free, 2= very mild, 3= discomforting, 4=tolerable, 5= very distressing, 6=intense, 7=very  
174 intense, 8=utterly horrible, 9=excruciating unbearable, 10=unimaginable unspeakable) [26],  
175 and participants were permitted to stop the test if pain levels were above bearable levels ( $\geq 4$ ).  
176 For the evaluation of isometric quadriceps torque using the LC, all positional measurements  
177 were recorded for replicability. Participants were seated on a standard chair, and the LC was



178 attached to a strap that was attached to a fixed surface behind the chair. The strap was  
179 fastened to the participant's ankle just above the lateral and medial malleoli where the  
180 distance between the lateral malleolus and the start of the ankle strap was measured to more  
181 accurately calculate the lever arm length in the evaluation of torque from the LC. The straps  
182 were adjusted until the knee was at 60° of flexion, measured with a digital goniometer. The  
183 participant completed the same warm-up and testing protocol as with the ID. The tests were  
184 repeated in the second session for the reliability analysis, and all procedures were repeated  
185 after 8-weeks. The conversion of raw kilogram values from the LC were converted to torque  
186 values using the following equation:

$$187 \text{ Torque (N.m) = LC 'force' (kg) } \times 9.81 \text{ (m.s}^{-2}\text{)} \times \text{LA distance (m)} \quad [\text{Equation 1}]$$

## 188 **Statistical analyses**

189 All data were evaluated for normality using the Shapiro-Wilk test, with deviations from  
190 normality being accepted at  $p < 0.05$ . All data are presented as mean  $\pm$  standard deviation  
191 (SD) unless otherwise stated. To determine the reliability of the LC, the intraclass correlation  
192 coefficient (ICC, two-way mixed effects, absolute agreement) between two measurements on  
193 separate occasions was used. The ICC values were interpreted as follows: poor:  $< 0.50$ ;  
194 moderate:  $0.50-0.75$ ; good:  $0.75-0.90$ ; excellent:  $>0.90$  [27]. Additional measures of  
195 reliability included: (i) typical error of measurement (TEM) (Equation 1), (ii) TEM%  
196 (Equation 3), and (iii) clinically meaningful difference (CMD) (Equation 4) [28]:

$$197 \text{ TEM} = \text{SD} \sqrt{(1-\text{ICC})} \quad [\text{Equation 2}]$$

$$198 \text{ TEM\%} = \text{TEM}/\text{mean} \times 100 \quad [\text{Equation 3}]$$

$$199 \text{ CMD} = \text{TEM} \times 1.96 \times \sqrt{2} \quad [\text{Equation 4}]$$

200 The TEM% scores were qualitatively interpreted as: excellent:<5%; good: 5-10%; poor:  
201 >10%. The CMD was used to evaluate the smallest detectable difference that would need to  
202 be overcome to conclude that a true change has occurred. A generalised linear model was  
203 used to evaluate mean differences between groups (2 levels: Con vs. PFP), and devices (2  
204 levels: ID vs. LC) where each participant was treated as a random effect in the model. Post-  
205 hoc analyses entailed the use of paired sample t-tests with a Holm correction to adjust for  
206 multiple comparisons. Standardised mean differences were calculated as Hedge's g, the  
207 magnitude of which was qualitatively interpreted as: trivial: <0.2; small: 0.2-0.6; moderate:  
208 0.6-1.2; and large: > 1.2 [29].

209 The concurrent validity was determined by using linear regression where the Pearson  
210 correlation coefficient (r), coefficient of determination ( $r^2$ ), standard error of the estimate  
211 (SEE), and the slope of the regression lines were evaluated. The magnitude of the correlation  
212 coefficients were qualitatively interpreted as follows: negligible: 0.00-0.10; weak: 0.10-0.39;  
213 moderate: 0.40-0.69; strong: 0.70-0.89; and very strong: 0.90-1.00 [30]. Bland-Altman  
214 analyses were used to determine the systematic bias between the LC (reference measure) and  
215 the ID (criterion measure) [31]. For both the regression and Bland-Altman analyses, the point  
216 estimates were evaluated for potential outliers using Cook's distance where potential outliers  
217 (PO) were flagged when the Cook's distance exceeded a given threshold calculated as:  $4/n$   
218 (where n is the number of observations) [32]. All statistical analyses were completed using R  
219 [33].

## 220 **Results**

221 The results pertaining to the relative ( $ICC_{3,1}$ ) and absolute (TEM) reliability as well as CMD  
222 of the LC are highlighted in Figure 1. Generally, the LC shows excellent relative and absolute  
223 reliability both at baseline and following the 8-week training interval. The ICC point

224 estimates exhibit fairly narrow confidence intervals for the control group, but marginally  
225 longer intervals for the PFP group, indicating greater variability in torque measures. In all  
226 instances the CMD exceeded the TEM which implies that the LC had sufficient sensitivity to  
227 measure a true change in isometric quadriceps strength (i.e. high signal-to-noise ratio).

228 **\*\*\* Figure 1 about here \*\*\***

229

230 Simple interaction effects for both between-group and within-group differences in mean  
231 torque values are shown in Figure 2. Mean differences between devices and group are shown  
232 in Figure 2A, whereas the standardised effect size with 95% CI and corresponding  
233 uncertainty density distribution for the point estimates are shown in Figure 2B.

234 **\*\*\* Figure 2 about here \*\*\***

235

236 For the concurrent validity, strong to very strong correlations were evident between the LC  
237 and ID torque measurements ( $r = 0.88-0.90$ ,  $SEE = 0.05-0.07$  Nm) (see Figure 3). Based on  
238 the slope analysis however, it is important to note that the torque values from the LC are  
239 typically lower than that obtained from the ID (slope =  $0.65-0.77$ ).

240 **\*\*\* Figure 3 about here \*\*\***

241

242

243 The results from the Bland-Altman analysis confirm the measurement bias between the LC  
244 and the ID whereby an average systematic bias of 16.3-28.8 Nm is evident in favour of the ID  
245 (see Figure 4).

246 **\*\*\* Figure 4 about here \*\*\***

247

248 Within-group differences in peak torque for each group and each device are shown in Figure  
249 5. Although no statistically significant differences are evident, it is important to emphasise  
250 the inter-individual variability highlighted by the colour-gradients which indicate the  
251 magnitude of participant-specific torque improvements in quadriceps strength between time  
252 points.

253

\*\*\* Figure 5 about here \*\*\*

254

## 255 Discussion

256 The present study yielded several novel findings. Firstly, the reliability and validity of a low-  
257 cost, commercially available LC were evaluated in comparison to a gold standard ID. The  
258 utility of such a device would have meaningful implications in rehabilitation services,  
259 especially in under-resourced practices. Secondly, we showed that the LC exhibited  
260 exceptional sensitivity for detection of clinically meaningful changes in isometric knee  
261 torque. Finally, we evaluated differences in isometric knee torque between a control and PFP  
262 group following an 8-week period showcasing substantial within-subject variability and  
263 constrained between-group differences.

264

### 265 Reliability

266

267 The LC showed excellent relative and absolute reliability in both the control and PFP groups  
268 (ICC>0.90 and TEM <10%) in all instances apart from the uninjured limb of the PFP group.  
269 The large ICC values obtained in the present study compare favourably with those from a  
270 separate study where belt-stabilisation was used with an HHD (ICC = 0.62 to 0.96) [34].

271 Similarly, the TEM% scores of the present study were substantially better those reported by  
272 Martins *et al.* [34] who recorded TEM% of 12% for a belt-stabilised HHD and Chamorro *et*  
273 *al.* [18] who yielded values ranging between 4-15% for measurements without stabilisation,  
274 both for knee extension strength. The ICC point estimates showed narrow confidence  
275 intervals in which the upper and lower limits fell within the margin defined as *excellent* in  
276 most instances. However, longer confidence intervals are observed in the PFP group than in  
277 the control group at the initial set of testing, indicating that the PFP group had greater  
278 variability in torque measures in the injured as well as the uninjured limbs compared to the  
279 control group. Torque variability for knee extensors is present in other knee injuries as well  
280 [35], and in this case the variability may be accounted for in the PFP group on the basis that  
281 PFP patients tend to have impaired quadriceps function often ascribed to impaired vastus  
282 medialis oblique firing [36,37] and reduced eccentric control [38]. However, the variability  
283 improved notably in the uninjured limb following the intervention period, possibly due to  
284 enhanced motor learning which tends to occur in injured populations [39].

285 It should be noted that the reliability of the LC, as with any device, is dependent on the set-up  
286 and therefor clinicians should be vigilant in following the set-up instructions set out in the  
287 procedures section and keep the set-up consistent when doing testing and re-testing in clinical  
288 settings, to ensure reliable results. Belt-stabilisation seems to be improving reliability in LC-  
289 based devices and are recommended to use by practitioners [25,34].

## 290 **Clinically meaningful difference**

291

292 The CMD within the present study ranged between 12-45 Nm, implying that the differences  
293 in torque production by the quadriceps musculature between limbs or over the course of  
294 rehabilitation can be detected with the LC. Previous research has shown that an HHD

295 exhibited comparable, although marginally lower CMD values of 17-27 Nm [40], with the ID  
296 yielding the lowest and most consistent CMD of approximately 27 Nm [41].

## 297 **Group differences in torque production**

298

299 An intriguing finding of the present study relates to the simple interaction effects regarding  
300 within-group (e.g., ID vs. LC) and between-group (e.g., Control vs. PFP) differences for  
301 quadriceps torque (see Figure 2A, 2B). The mean differences in torque production between  
302 devices and groups are classified as trivial-to-moderate, and typically range -2.28 Nm to  
303 27.83 Nm ( $p = 0.113-0.999$ ). In most instances the point estimates of the mean difference  
304 were not significantly different from zero largely due to the longer confidence intervals. Such  
305 a result generally implies, at least in principle, similarities in torque production between  
306 groups and between devices. The standardised effect sizes (together with the 95% CI) show  
307 that, despite a lack of statistical significance, the magnitude of the mean differences are likely  
308 to be meaningful, especially for the PFP group which consistently yielded higher torque  
309 values on the ID compared to the LC ( $M_{diff} = 43.53$  Nm,  $p < 0.001$ ). The interaction effects  
310 should be of interest to practitioners on the basis that the interchangeability of device  
311 measurements must be considered when interpreting results and making clinical inferences  
312 [41]. Given that, at least in some instances, practitioners might receive isokinetic results for a  
313 specific PFP patient or a referring clinician, these results should be verified when using the  
314 LC in the clinical setting so that more precise interpretations can be made when evaluating  
315 temporal changes in quadriceps strength. Moreover, the same device should be used by the  
316 same clinician to ensure adequate consistency in readings and insights of whether true  
317 changes have accrued.

## 318 **Validity**

319

320 The results of the preset study suggest that the concurrent validity of the LC, when compared  
321 to the ID, appears to be strong to very strong, although a few discrepancies are noteworthy.  
322 The torque values from the LC are typically lower than those from the ID, with the exception  
323 of torque values exceeding ~300 Nm (see Figure 3 and Figure 4). It is therefore reasonable to  
324 state that the LC might be of better use in injured and non-athletic populations that will most  
325 likely produce lower torque values compared to healthy, uninjured, or very athletic  
326 participants that might produce substantially higher torque values. The validity of the LC was  
327 greater than that of an HHD which yielded poorer scores even when stabilised with a belt,  
328 with correlations ranging between  $r = 0.3-0.8$  [18,42]. The LC is not a perfectly valid tool  
329 given the bias in torque readings and moderate-to-large limits of agreement (LoA) and is  
330 therefore unlikely to replace the ID for absolute values. More specifically, there is evidence  
331 of a systematic bias in favour of the ID (see Figure 4), confirming that the LC typically  
332 measures lower torques than the ID on average. These results correspond with Martins et al.  
333 [34], who showed that knee extension tested with a belt-stabilised HHD, exhibited similar  
334 mean difference in torque production, with bias towards the ID. The HHD also shows, on  
335 average, wide LoA values (33.59%,  $CI_{95\%}$  [23.91%, 43.26%]) for knee extension in other  
336 literature [18]. The wide LoA therefore support the notion that the LC is not valid enough to  
337 replace values obtained from the ID but given the excellent reliability would still be an  
338 exceptional tool for the evaluation of isometric quadriceps strength. It is also important to  
339 highlight that isokinetic norms should not be used to make clinical interpretations when  
340 comparing these to values derived from the LC. It would be important to develop  
341 independently generated normative data for LC-derived torque values to facilitate decision-  
342 making across different joints and population groups.

### 343 **Temporal changes in torque**

344

345 Finally, although the mean peak torque values did not change significantly from the baseline  
346 to follow-up testing ( $M_{diff} = 0.77-14.63$  Nm,  $p = 0.531-0.969$ ), it should be noted that there  
347 were substantial individual improvements both within- and between groups (see Figure 5).  
348 The variability in individual responses underscores the potential inadequacy of the  
349 intervention programmes followed by the groups which were beyond the control of the  
350 current study. Whether more focused and targeted interventions or prolonged rehabilitation  
351 timelines (> 8 weeks) would elicit more favourable outcomes would require further research,  
352 especially in those with PFP. It should however be noted that there were consistencies in the  
353 measured values between devices across time, implying that the LC indeed has exceptional  
354 utility as a measurement tool for evaluating isometric quadriceps strength.

## 355 **Limitations**

356

357 Given the strengths of the present study, it is also important to underscore some limitations.  
358 Firstly, participants could follow any rehabilitation program at any practitioner of their  
359 choice, and therefore could not control the details associated with targeted quadriceps  
360 strengthening as this was beyond the scope of the present study. Secondly, the ratios of males  
361 versus females in the groups differed, although consistency within groups were more  
362 important for this study than consistency between groups. Thirdly, participants were not  
363 divided into different groups according to the magnitude of quadriceps strength deficits. This  
364 latter point could provide more nuanced information on strength improvements over time in  
365 future studies.

## 366 **Conclusions**

367 The objectives of this study were to determine the validity and reliability of a low-cost LC for  
368 use in a clinical setting for measuring knee extension strength in those with PFP. The LC  
369 exhibited excellent reliability, and was deemed sensitive enough to detect clinically



370 meaningful differences over time in both healthy individuals as well as those with PFP. The  
371 reliability of the LC is dependent on the set-up of the individual, and therefore practitioners  
372 should take care to complete the set-up as described and ensure consistency across every  
373 testing session. The LC lacks validity and is therefore unlikely to be an adequate surrogate for  
374 isokinetic dynamometry. However, given the trade-off in (i) costing associated with the ID  
375 and the LC, (ii) the importance of evaluating and tracking changes in knee extension strength  
376 in those with PFP, and (iii) the excellent reliability of the LC, the utility of the LC as a viable  
377 assessment tool is advocated, especially in resource-restricted settings.

## 378 **References**

- 379 1. Crossley KM, Stefanik JJ, Selfe J, Collins NJ, Davis IS, Powers CM, et al. 2016 Patellofemoral pain  
380 consensus statement from the 4th International Patellofemoral Pain Research Retreat,  
381 Manchester. Part 1: Terminology, definitions, clinical examination, natural history, patellofemoral  
382 osteoarthritis and patient-reported outcome measures. *Br J Sports Med* [Internet].  
383 2016;50(14):839–43. Available from: <https://bjsm.bmj.com/content/50/14/831>
- 384 2. Hall R, Barber Foss K, Hewett TE, Myer GD. Sport specialization's association with an increased  
385 risk of developing anterior knee pain in adolescent female athletes. *J Sport Rehabil*.  
386 2015;24(1):31–5.
- 387 3. Dey P, Callaghan M, Cook N, Sephton R, Sutton C, Hough E, et al. A questionnaire to identify  
388 patellofemoral pain in the community: An exploration of measurement properties. *BMC*  
389 *Musculoskeletal Disorders*. 2016;17(1):e237.
- 390 4. Roush JR, Bay RC. Prevalence of anterior knee pain in 18-35 year-old females. *J Sports Phys Ther*.  
391 2012;7(4):396–401.
- 392 5. Kunene SH, Ramklass S, Taukobong NP. Anterior knee pain and its intrinsic risk factors among  
393 runners in under-resourced communities in Ekurhuleni, Gauteng. *S Afr J Physiother*.  
394 2018;74(1):e452.
- 395 6. Coburn SL, Barton CJ, Filbay SR, Hart HF, Rathleff MS, Crossley KM. Quality of life in individuals  
396 with patellofemoral pain: A systematic review including meta-analysis. *Physical Therapy in Sport*  
397 [Internet]. 2018;33:96–108. Available from: <https://doi.org/10.1016/j.ptsp.2018.06.006>
- 398 7. Collins NJ, Bierma-Zeinstra SMA, Crossley KM, Van Linschoten RL, Vicenzino B, Van Middelkoop  
399 M. Prognostic factors for patellofemoral pain: a multicentre observational analysis. *Br J Sports*  
400 *Med* [Internet]. 2013;47(4):227–33. Available from:  
401 <https://pubmed.ncbi.nlm.nih.gov/23242955/>
- 402 8. Reijnders L, Van de Groes SAW. The quality of life of patients with patellofemoral pain - a  
403 systematic review. *Acta Orthop Belg*. 2020;86(4):678–87.
- 404 9. Tan SS, van Linschoten RL, van Middelkoop M, Koes BW, Bierma-Zeinstra SM, Koopmanschap  
405 MA, et al. Cost-utility of exercise therapy in adolescents and young adults suffering from the  
406 patellofemoral pain syndrome. *Scand J Med Sci Sports*. 2010;20(4):568–79.
- 407 10. Nimon G, Murray D, Sandow M, Goodfellow J. Natural History of Anterior Knee Pain: A 14- to 20-  
408 Year Follow-up of Nonoperative Management [Abstract]. *Journal of Pediatric Orthopaedics*  
409 [Internet]. 1998 [cited 2023 Apr 17];18(1):118–22. Available from:  
410 [https://journals.lww.com/pedorthopaedics/Abstract/1998/01000/Natural\\_History\\_of\\_Anterior\\_Knee\\_Pain\\_\\_A\\_14\\_\\_to.21.aspx](https://journals.lww.com/pedorthopaedics/Abstract/1998/01000/Natural_History_of_Anterior_Knee_Pain__A_14__to.21.aspx)  
411
- 412 11. Thomas MJ, Wood L, Selfe J, Peat G. Anterior knee pain in younger adults as a precursor to  
413 subsequent patellofemoral osteoarthritis: a systematic review. *BMC Musculoskelet Disord*  
414 [Internet]. 2010 [cited 2023 Apr 17];11(1):e201. Available from:  
415 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2944218/>

- 416 12. Van Middelkoop M, Bennell KL, Callaghan MJ, Collins NJ, Conaghan PG, Crossley KM, et al.  
417 International patellofemoral osteoarthritis consortium: Consensus statement on the diagnosis,  
418 burden, outcome measures, prognosis, risk factors and treatment. *Semin Arthritis Rheum*.  
419 2018;47(5):666–75.
- 420 13. Boling MC, Padua DA, Marshall SW, Guskiewicz K, Pyne S, Beutler A. A Prospective Investigation  
421 of Biomechanical Risk Factors for Patellofemoral Pain Syndrome The Joint Undertaking to  
422 Monitor and Prevent ACL Injury (JUMP-ACL) Cohort. *Am J Sports Med* [Internet].  
423 2009;37(11):2108–16. Available from: <http://ajs.sagepub.com/supplemental/>
- 424 14. Amin S, Baker K, Niu J, Clancy M, Goggins J, Guermazi A, et al. Quadriceps Strength and the Risk  
425 of Cartilage Loss and Symptom Progression in Knee Osteoarthritis. *Arthritis Rheum* [Internet].  
426 2009 [cited 2023 Apr 17];60(1):189–98. Available from:  
427 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653642/>
- 428 15. Willy RW, Hoggund LT, Barton CJ, Bolgla LA, Scalzitti DA, Logerstedt DS, et al. Patellofemoral pain.  
429 *J Orthop Sports Phys Ther*. 2019;49(9):CPG1–95.
- 430 16. Janik F, Toulotte C, Seichepine AL, Masquelier B, Barbier F, Fabre C. Isometric Strength Database  
431 for Muscle Maximal Voluntary Endurance Field Tests: Normative Data. *Sports Med* [Internet].  
432 2021;7(1):e47. Available from: [/pmc/articles/PMC8273050/](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC8273050/)
- 433 17. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held Dynamometry Correlation With the Gold  
434 Standard Isokinetic Dynamometry: A Systematic Review. *PM&R* [Internet]. 2011;3(5):472–9.  
435 Available from: [https://onlinelibrary-wiley-](https://onlinelibrary-wiley-com.nwulib.nwu.ac.za/doi/full/10.1016/j.pmrj.2010.10.025)  
436 [com.nwulib.nwu.ac.za/doi/full/10.1016/j.pmrj.2010.10.025](https://onlinelibrary-wiley-com.nwulib.nwu.ac.za/doi/full/10.1016/j.pmrj.2010.10.025)
- 437 18. Chamorro C, Armijo-Olivo S, De La Fuente C, Fuentes J, Javier Chiroso L. Absolute reliability and  
438 concurrent validity of hand held dynamometry and isokinetic dynamometry in the hip, knee and  
439 ankle joint: Systematic review and meta-analysis. *Open Med* [Internet]. 2017;12(1):359–75.  
440 Available from: [/pmc/articles/PMC5651404/](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC5651404/)
- 441 19. Deones VL, Wiley SC, Worrell T. Assessment of quadriceps muscle performance by a hand-held  
442 dynamometer and an isokinetic dynamometer. *JOSPT* [Internet]. 1994;20(6):296–301. Available  
443 from: <https://pubmed.ncbi.nlm.nih.gov/7849749/>
- 444 20. Christensen LB, Johnson RB, Turner LA. Research methods, design and analysis. 12th ed. Vol. 18.  
445 Harlow: Pearson Education Limited; 2015.
- 446 21. Faul F, Erdfelder E, Lang A, Buchner A. Statistics textbooks in the social. *Behavior Research*  
447 *Methods* [Internet]. 2007;39(2):175–91. Available from:  
448 <https://link.springer.com/content/pdf/10.3758/BF03193146.pdf>
- 449 22. Howe TE, Dawson LJ, Syme G, Duncan L, Reid J. Systematic review Evaluation of outcome  
450 measures for use in clinical practice for adults with musculoskeletal conditions of the knee: A  
451 systematic review. *Man Ther*. 2012;17(2):100–18.
- 452 23. Watson CJ, Propps M, Ratner J, Zeigler DL, Horton P, Smith SS. Reliability and responsiveness of  
453 the lower extremity functional scale and the anterior knee pain scale in patients with anterior  
454 knee pain. *J Orthop Sports Phys Ther* [Internet]. 2005;35(3):136–46. Available from:  
455 <https://pubmed.ncbi.nlm.nih.gov/15839307/>

- 456 24. Hasan S. Quadriceps Femoris Strength Training: effect of Neuromuscular Electrical Stimulation Vs  
457 Isometric Exercise in Osteoarthritis of Knee. *Indian J Physiother Occup Therapy*. 2015;9(3):129–  
458 34.
- 459 25. McNamara S, Hevia EG, St. Louis R, Cho W, Lee S, Moyne M, et al. Isometric quadriceps strength  
460 test device to improve the reliability of handheld dynamometry in patient with anterior cruciate  
461 ligament injury. In Minneapolis, MN, USA: ResearchGate; 2018. Available from:  
462 [http://asmedigitalcollection.asme.org/BIOMED/proceedings-](http://asmedigitalcollection.asme.org/BIOMED/proceedings-pdf/DMD2018/40789/V001T09A001/2788227/v001t09a001-dmd2018-6848.pdf)  
463 [pdf/DMD2018/40789/V001T09A001/2788227/v001t09a001-dmd2018-6848.pdf](http://asmedigitalcollection.asme.org/BIOMED/proceedings-pdf/DMD2018/40789/V001T09A001/2788227/v001t09a001-dmd2018-6848.pdf)
- 464 26. Aby F. Serotonergic neurons of the nucleus raphe Magnus in the control of nociceptive  
465 transmission in the dorsal horn of the spinal cord: an optogenetic study in different  
466 pathophysiological contexts. [Nouvelle-Aquitaine, France]: University of Bordeaux; 2019.
- 467 27. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for  
468 Reliability Research. *J Chiropr Med [Internet]*. 2016;15(2):155–155. Available from:  
469 </pmc/articles/PMC4913118/>
- 470 28. Schwenk M, Gogulla S, Englert S, Czempik A, Hauer K. Test–retest reliability and minimal  
471 detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric  
472 patients. *Physiol Meas [Internet]*. 2012 Oct [cited 2024 Jan 8];33(11):1931. Available from:  
473 <https://dx.doi.org/10.1088/0967-3334/33/11/1931>
- 474 29. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports  
475 medicine and exercise science. *Med Sci Sports Exerc*. 2009 Jan;41(1):3–13.
- 476 30. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation.  
477 *Anesthesia & Analgesia [Internet]*. 2018 May [cited 2024 Jan 8];126(5):1763. Available from:  
478 [https://journals.lww.com/anesthesia-](https://journals.lww.com/anesthesia-analgesia/fulltext/2018/05000/correlation_coefficients__appropriate_use_and.50.aspx)  
479 [analgesia/fulltext/2018/05000/correlation\\_coefficients\\_\\_appropriate\\_use\\_and.50.aspx](https://journals.lww.com/anesthesia-analgesia/fulltext/2018/05000/correlation_coefficients__appropriate_use_and.50.aspx)
- 480 31. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb) [Internet]*. 2015 Jun 5  
481 [cited 2023 Dec 19];25(2):141–51. Available from:  
482 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4470095/>
- 483 32. Hebbali A. [rsquaredacademy/olsrr](https://github.com/rsquaredacademy/olsrr) [Internet]. Rsquared Academy; 2023 [cited 2024 Jan 17].  
484 Available from: <https://github.com/rsquaredacademy/olsrr>
- 485 33. R Core Team. R: The R Project for Statistical Computing [Internet]. Vienna, Austria: R Foundation  
486 for Statistical Computing; [cited 2024 Jan 8]. Available from: <https://www.r-project.org/>
- 487 34. Martins J, Rodrigues Da Silva J, Rodrigues Barbosa Da Silva; D ´ Eborá Bevilaqua-Grossi M.  
488 Reliability and Validity of the Belt-Stabilized Handheld Dynamometer in Hip-and Knee-Strength  
489 Tests. *J Athl Train [Internet]*. 2017;52(9):809–19. Available from: [www.natajournals.org](http://www.natajournals.org)
- 490 35. Goetschius J, Hart JM. Knee-Extension Torque Variability and Subjective Knee Function in  
491 Patients with a History of Anterior Cruciate Ligament Reconstruction. *J Athl Train*. 2016  
492 Jan;51(1):22–7.
- 493 36. Van Tiggelen D, Cowan S, Coorevits P, Duvigneaud N, Witvrouw E. Delayed vastus medialis  
494 obliquus to vastus lateralis onset timing contributes to the development of patellofemoral pain  
495 in previously healthy men: a prospective study [Abstract]. *Am J Sports Med*. 2009;37(6):1099–  
496 105.

- 497 37. Pal S, Draper CE, Fredericson M, Gold GE, Delp SL, Beaupre GS, et al. Patellar maltracking  
498 correlates with vastus medialis activation delay in patellofemoral pain patients. *Am J Sports*  
499 *Med.* 2011;39(3):590–8.
- 500 38. Eapen C, Nayak CD, Pazhyaottyil Zulfeequer C. Effect of Eccentric Isotonic Quadriceps Muscle  
501 Exercises on Patellofemoral Pain Syndrome: An Exploratory Pilot Study. *Asian J Sports Med*  
502 [Internet]. 2011 [cited 2023 Jun 3];2(4):227–34. Available from:  
503 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3289225/>
- 504 39. Rodrigues-da-Silva JM, de Rezende MU, Spada TC, da Silva Francisco L, Greve JMD, Ciolac EG.  
505 Effects of Motor Learning on Clinical Isokinetic Test Performance in Knee Osteoarthritis Patients.  
506 *Clinics (Sao Paulo)* [Internet]. 2017 [cited 2024 Jan 12];72(4):202–6. Available from:  
507 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5401619/>
- 508 40. Knols RH, Aufdemkampe G, de Bruin ED, Uebelhart D, Aaronson NK. Hand-held dynamometry in  
509 patients with haematological malignancies: Measurement error in the clinical assessment of  
510 knee extension strength. *BMC Musculoskelet Disord* [Internet]. 2009 Mar 9 [cited 2024 Jan  
511 12];10:31. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2662793/>
- 512 41. de Araujo Ribeiro Alvares JB, Rodrigues R, de Azevedo Franke R, da Silva BGC, Pinto RS, Vaz MA,  
513 et al. Inter-machine reliability of the Biodex and Cybex isokinetic dynamometers for knee  
514 flexor/extensor isometric, concentric and eccentric tests. *Physical Therapy in Sport* [Internet].  
515 2015 Feb 1 [cited 2024 Jan 17];16(1):59–65. Available from:  
516 <https://www.sciencedirect.com/science/article/pii/S1466853X14000339>
- 517 42. Hirano M, Katoh M, Gomi M, Arai S. Validity and reliability of isometric knee extension muscle  
518 strength measurements using a belt-stabilized hand-held dynamometer: a comparison with the  
519 measurement using an isokinetic dynamometer in a sitting posture. *Journal of Physical Therapy*  
520 *Science.* 2020;32(2):120–4.
- 521

## 522 **Figure Captions**

523 Figure 1: Between-group and within-group intra-class correlation coefficients. The ICC  
524 values are shown with their 95% confidence intervals. Above each point estimate are shown  
525 the precise ICC values, typical error of measurement (TEM%) and clinically meaningful  
526 difference (CMD) to provide greater context for the reliability score. Vertical dotted lines  
527 represent thresholds for excellent (ICC > 0.90) and good (ICC: 0.70-0.90). Note: ND = non-  
528 dominant; Dom = dominant; Con = control group; PFP = patellofemoral pain group.

529

530 Figure 2: Absolute and relative within-group and between-group mean differences. Panel A:  
531 Mean differences with 95% CI; panel B: standardised mean effect size (Hedge's  $g$ ) with 90%  
532 CI (thick black line) and 95% CI (thin black line) as well as density estimates to highlight the  
533 uncertainty in the point estimate. Note:  $M_{diff}$  = mean difference;  $g$  = Hedge's  $g$  effect size;  
534 Con = control group; PFP = patellofemoral pain group.

535

536 Figure 3: Linear regression between load-cell and Cybex for isometric knee extension torque.  
537 PO = potential outlier; SEE = standard error of the estimate; ND = non-dominant; Dom =  
538 dominant

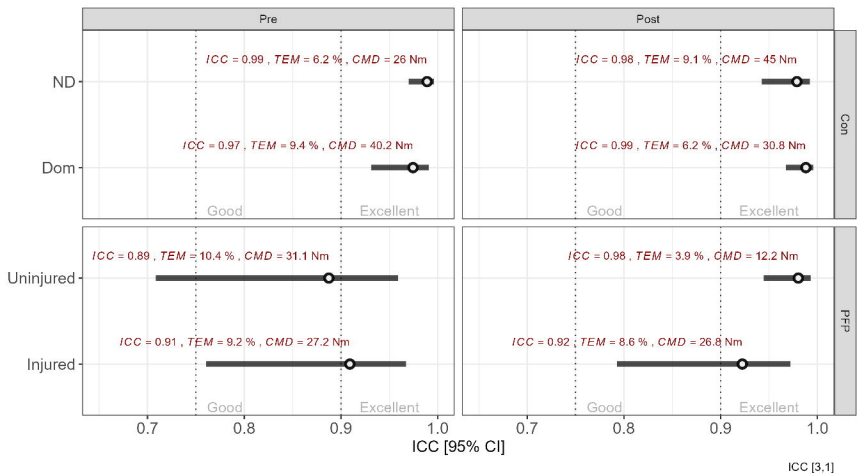
539

540 Figure 4: Bland-Altman plots for the bias (95% CI) between instruments and the limits of  
541 agreement. Note: UL = upper limit; LL = lower limit; ND = non-dominant; Dom = dominant;  
542 PO = potential outlier

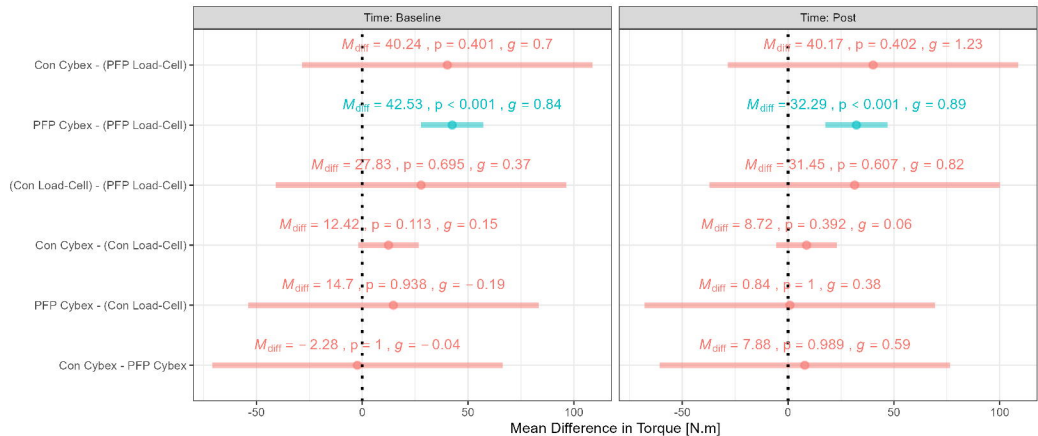
543

544 Figure 5: Within-group differences in peak torque. Pre-post differences are shown for the  
545 control group as evaluated by the LC (panel A) and ID (panel B). Pre-post differences are  
546 shown for the PFP group as evaluated by the LC (panel C) and ID (panel D). Line segments  
547 are coloured based on the percentage difference (Post-Pre/Pre $\times$ 100) where improvements are  
548 shifted towards orange, and decrements are shifted towards green. The magnitude of the  
549 difference is captured by the intensity of the colour gradient (larger differences are darker,  
550 smaller differences are lighter). Black lines indicate the group mean for a given time point.  
551 Note:  $M_{diff}$  = mean difference;  $g_{Hedges}$  = Hedge's  $g$ ; Con = control group; PFP =  
552 patellofemoral pain group; LC = Load Cell

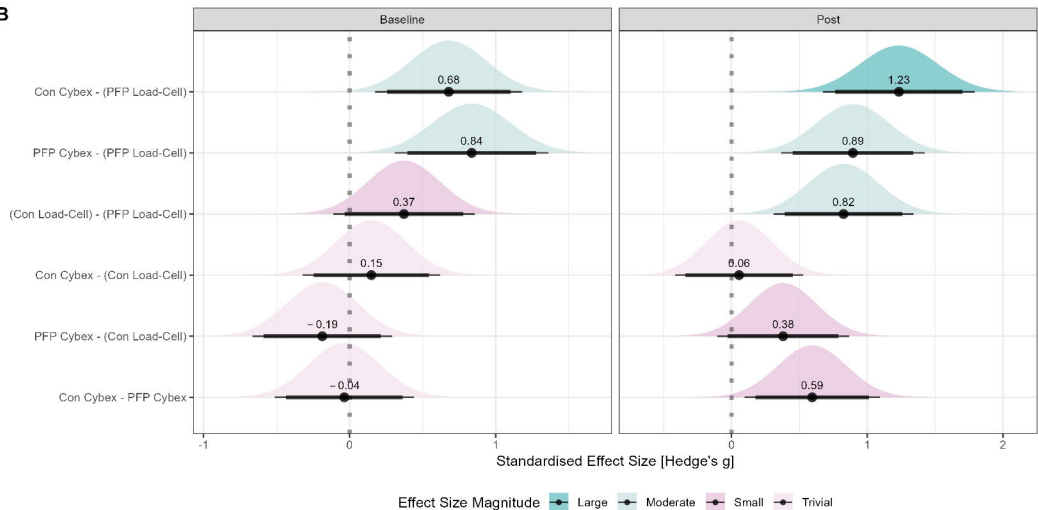
553



Significant ● No ● Yes



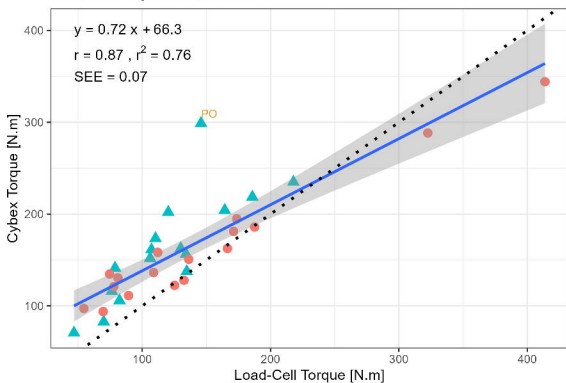
**B**



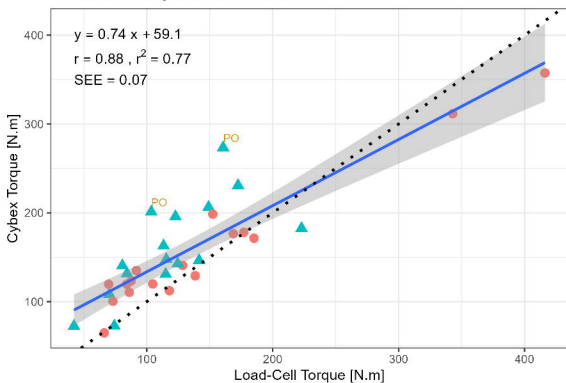


Group ● Con ▲ PFP

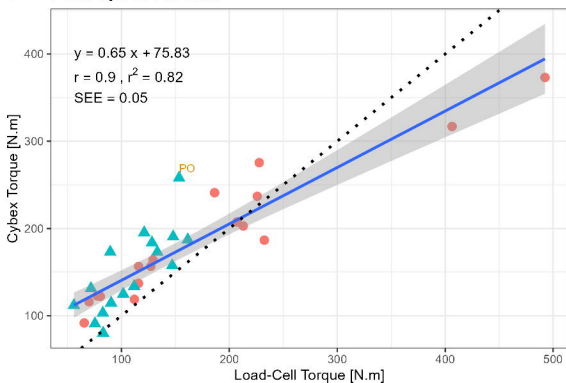
**A** Baseline: Injured-ND Limb



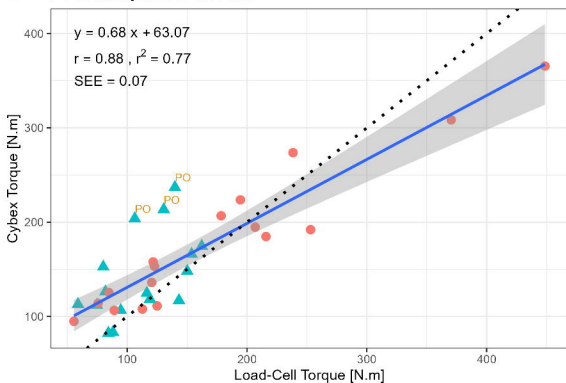
**B** Baseline: Uninjured-Dom Limb



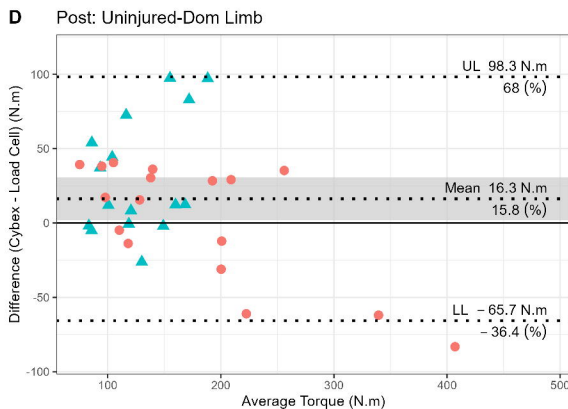
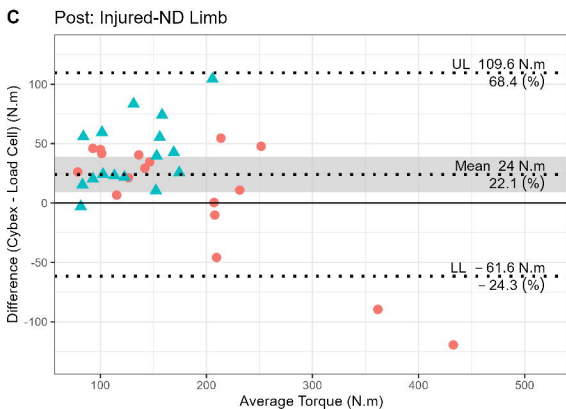
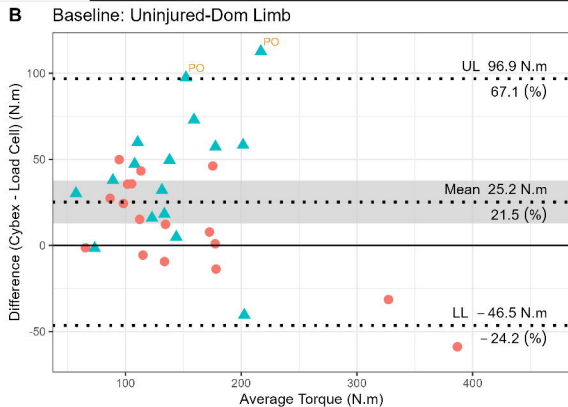
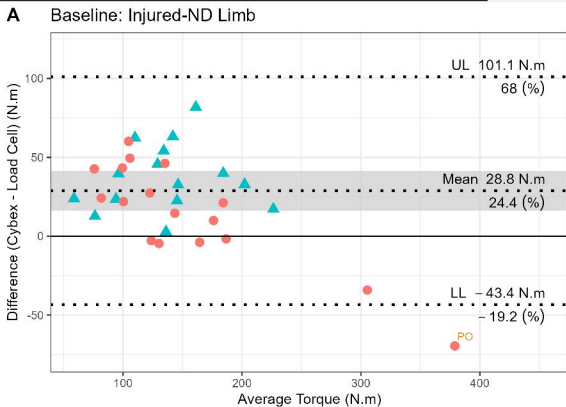
**C** Post: Injured-ND Limb



**D** Post: Uninjured-Dom Limb



Group ● Con ▲ PFP



% Difference <-30 -20 -10 0 10 20 >30

