

RESEARCH ARTICLE

Assessing quadriceps strength in patellofemoral pain patients: A study on the reliability and validity of a low-cost strain-gauge for clinical practice

Germari Deysel, Mariaan van Aswegen, Mark Kramer *

Physical Activity, Sport, and Recreation (PhasRec) Research Focus Area, North-West University, Potchefstroom, North West Province, South Africa

* mark.kramer@nwu.ac.za

Abstract

Background

Patellofemoral pain (PFP) is a common knee complaint affecting diverse populations both acutely and chronically. Quadriceps muscle weakness is one possible aetiology, but current devices for measuring muscle strength (isokinetic dynamometer [ID] and hand-held dynamometers [HHD]) are frequently too expensive (e.g., ID) or lack reliability (e.g., HHD) for practitioners, especially in under-resourced settings. There is a need to evaluate a low-cost device to manage rehabilitation of people with PFP.

Methods

Isometric quadriceps strength of participants aged 18–35 years (total [n = 33], control group [n = 17] and PFP group [n = 16]) were evaluated on an isokinetic dynamometer and a commercially available strain gauge at baseline and after an 8-week non-standardised intervention.

Results

The strain gauge showed high absolute and relative reliability (intraclass correlation coefficient = 0.89–0.99; typical error of measurement = 3.9–10.4%). Clinically meaningful difference scores (12.2–45 Nm) were greater than the typical error of measurement, implying sufficient sensitivity of the strain gauge to measure true changes in isometric quadricep strength. Strong to very strong correlations were evident between the strain gauge and isokinetic dynamometer torque measurements ($r = 0.88–0.90$, $SEE = 0.05–0.07$ Nm), but slope values ($\beta = 0.65–0.77$) indicated that torque from the strain gauge was lower than that obtained from the isokinetic dynamometer. An average systematic bias of 16.3–28.8 Nm was evident in favour of the isokinetic dynamometer, with no statistically significant between-group differences apparent between baseline and follow-up testing.

OPEN ACCESS

Citation: Deysel G, van Aswegen M, Kramer M (2024) Assessing quadriceps strength in patellofemoral pain patients: A study on the reliability and validity of a low-cost strain-gauge for clinical practice. PLoS ONE 19(5): e0298570. <https://doi.org/10.1371/journal.pone.0298570>

Editor: Theodoros M. Bampouras, Liverpool John Moores University, UNITED KINGDOM

Received: January 28, 2024

Accepted: May 9, 2024

Published: May 28, 2024

Copyright: © 2024 Deysel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data underlying the results presented in the study are available from the Harvard Dataverse under the following URL: <https://doi.org/10.7910/DVN/N1SSPT>.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

The present commercially available strain gauge is reliable and sensitive enough to detect clinically meaningful differences in quadriceps strength of both healthy individuals and those with PFP. However, the strain gauge lacks validity and therefore cannot replace isokinetic dynamometry. Given the low cost and excellent reliability, the strain gauge can be a valuable tool to assess quadriceps muscle deficits and track rehabilitation progress in people with PFP.

Introduction

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

Given that people with PFP can be divided into different subgroups dependent on the individual aetiologies [15], it is important to determine whether quadriceps weakness is part of the cause for PFP in the individual [15]. Moreover, it is also essential to track whether progress is made when conducting an exercise program to improve quadriceps strength such that more sound and robust clinical decisions can be made [15]. Presently, the ‘gold standard’ instrument for the evaluation of quadriceps strength is the isokinetic dynamometer (ID) [16, 17], which is beyond the affordability of most clinicians. Although more reasonable alternatives to the ID exists, such as a handheld dynamometer (HHD) [17], the reliability of an HHD depends on the strength and aptitude of the practitioner [18, 19]. The latter is especially true when assessing stronger individuals or muscle strength around larger joints such as the knee joint [17] which may have important clinical implications for both patients and practitioners. Furthermore, HHDs are still relatively expensive when considering starting a rehabilitation practice, especially in under-resourced communities and clinical settings. In some instances, custom strain gauges have also been investigated [20] which show promise, but are presently not commercially available, at least in a general sense.

Therefore, the purpose of this study was to: (i) establish the validity and reliability of a low-cost commercially available strain gauge (SG) for the assessment of isometric quadriceps muscle strength in people with PFP, (ii) evaluate the clinically meaningful difference (CMD) needed to enhance practitioner-based decision-making, and (iii) assess differences in isometric quadriceps strength with the use of a low-cost device between those with PFP and healthy control following an 8-week intervention.

Materials and methods

Study design

This study used a repeated measures mixed study design consisting of a combination of between and within subject factors [21]. The between-subjects independent variable was

group allocation (i.e., control group and PFP group), where group assignment was non-randomised (due to presence/absence of PFP). Within-subjects independent variables included time (repeated measures) and extremity (injured [or non-dominant] and uninjured [or dominant] limb).

Participants

A total of 35 participants volunteered for the study, of which 17 were part of the control group (female [n = 10] and male [n = 7]) and 18 were part of the PFP group (female [n = 15] and male [n = 3]). Given the requirements of the study design, a minimum total sample size of 24 was calculated based on a repeated measures analysis of variance (ANOVA) design that incorporated (i) a within-between interaction, (ii) a moderate effect size ($f = 0.25$), (iii) a type-I error rate of 5% ($\alpha = 0.05$), (iv) a type-II error rate of 20% ($\beta = 0.20$), (v) 2 groups (control [n = 17] and PFP [n = 17]), (vi) 2 repeated measurements, (vii) an anticipated dropout of 10%, and (viii) a minimum expected correlation of 0.50 among repeated measurements [22]. Accounting for a potential drop-out of 20%, the minimum sample size for adequate statistical power was 29 participants. Two participants were lost to follow-up, both in the PFP group (female [n = 1] and male [n = 1]), resulting in a final sample size of 33 participants.

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

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

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

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

Instruments

A general demographic questionnaire was used to obtain the participant's contact details, e-mail address, age, sex, involved limb, dominant limb, previous injuries and PFP symptoms. The anterior knee pain scale (AKPS), also known as the Kujula patellofemoral scoring system, was completed electronically to capture knee-related pathologies where the total score out of 100 was captured. The AKPS was the chosen questionnaire, as it was developed specifically for evaluation of pain and disability in individuals with PFP [23], showed high test-retest reliability (intraclass correlation coefficient [ICC] = 0.95), and exhibited moderate responsiveness to

clinical change, which implies that the score will likely reflect meaningful changes in a patient's condition over time [24]. Body mass and stature were measured to calculate body mass index (BMI) [25]. Body mass was measured with an electronic scale (Seca 874, Seca, Germany) to the nearest 0.1 kg and stature with a portable stadiometer (Holtain Ltd., U.K.) to the nearest 0.01 m. A cycle ergometer (Wattbike Pro, Wattbike Ltd, Nottingham, UK) was used for a ten-minute warm-up before the testing on the ID for optimal muscle performance and reduced risk of injury. The ID (Cybex II, CSMi, Stoughton, MA, USA) was used as the gold standard for evaluating isometric muscle strength and a SG (Crane & Hanging Scale, Micro Mini CS300, Border Scales & Labels) served as the low-cost alternative to measure isometric quadriceps strength. Strength was evaluated with the knee joint at 60° of knee flexion, where peak force of the quadriceps is usually generated [26]. A digital goniometer (EasyAngle, Meloq AB, Stockholm, Sweden) was used to measure the knee angle in all instances to ensure true validity and replicability.

Procedures

The demographic questionnaire was completed first, and only at the baseline testing. Thereafter the participants completed the AKPS questionnaire verbally, where the score was calculated out of 100 and recorded electronically. Stature and body mass were measured during barefoot standing and with minimal clothing. Lower leg length was measured from the lateral condyle of the femur to the lateral malleolus of the tibia to use as the lever length in the formula to convert force to torque for the SC measurements. Prior to testing participants completed a ten-minute warm-up on a cycle ergometer at lowest resistance and a comfortable speed (rating of perceived exertion [RPE] < 2 on the modified Borg scale). Participants completed baseline testing in two sessions separated by 24 hours, and then repeated the same testing eight weeks later. During the first session, isometric strength of the quadriceps muscles was measured first with the ID, and thereafter with the SG. During the second session, only the testing on the SG was completed for the reliability analysis. Reliability analysis was done both at baseline testing and at follow-up testing conducted after 8 weeks to ensure a stronger reliability score.

For the evaluation of isometric quadriceps strength on the ID, all variables related to the set-up were recorded to replicate the exact position for follow-up testing (Fig 1A). The dominant limb (defined as the preferred kicking limb) in the control group and the uninjured limb in the PFP group were measured first. The participants were placed in a seated position, stabilized with upper body straps and an upper leg strap just above the knee joint. The knee was positioned at a 60° angle, using the machine angle provided by the dynamometer. The lateral femoral epicondyle was aligned with the axis of rotation of the dynamometer, and the resistance pad was positioned anterior to the distal tibia just superior to the lateral and medial malleoli. A gravity correction was performed to account for any potential additional torque induced on the attachment. The participant completed a warm-up round consisting of three repetitions, with the instruction to do one repetition at approximately 25%, 50% and 75% of perceived maximum effort before commencing of the test to assist with familiarisation of the test. A 10 second rest period was allotted between trial repetitions. A one-minute rest period was granted after the warm-up, after which the participant completed repetitions of five second maximal extension contraction with ten seconds of rest between repetitions. Verbal encouragement was given throughout the whole procedure. Pain levels were monitored by asking the participant to rate pain on a scale of 1–10 throughout the procedure (where 1 = pain free, 2 = very mild, 3 = discomforting, 4 = tolerable, 5 = very distressing, 6 = intense, 7 = very intense, 8 = utterly horrible, 9 = excruciating unbearable, 10 = unimaginable unspeakable) [27], and participants were permitted to stop the test if pain levels were above bearable



Fig 1. Experimental setup for the isokinetic dynamometer (panel A) and the strain gauge (panel B) for the right leg of a representative participant. The inset of panel C is a close-up of the commercially available strain gauge used in the study.

<https://doi.org/10.1371/journal.pone.0298570.g001>

levels (≥ 4). For the evaluation of isometric quadriceps torque using the SG, all positional measurements were recorded for replicability. Participants were seated on a standard chair, and the SG was attached to a strap that was attached to a fixed surface behind the chair (Fig 1B). The strap was fastened to the participant's ankle just above the lateral and medial malleoli where the distance between the lateral malleolus and the start of the ankle strap was measured to more accurately calculate the lever arm length in the evaluation of torque from the SG. The straps were adjusted until the knee was at 60° of flexion, measured with a digital goniometer. The participant completed the same warm-up and testing protocol as with the ID. The tests were repeated in the second session for the reliability analysis, and all procedures were repeated after 8-weeks. The conversion of raw kilogram values from the SG were converted to torque values using the following equation:

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

Statistical analyses

All data were evaluated for normality using the Shapiro-Wilk test, with deviations from normality being accepted at $p < 0.05$. All data are presented as mean \pm standard deviation (SD) unless otherwise stated. To determine the reliability of the SG, the intraclass correlation coefficient (ICC, two-way mixed effects, absolute agreement) between two measurements on separate occasions was used. The ICC values were interpreted as follows: poor: < 0.50 ; moderate: $0.50\text{--}0.75$; good: $0.75\text{--}0.90$; excellent: >0.90 [28]. Additional measures of reliability included:

(i) typical error of measurement (TEM) (Eq 1), (ii) TEM% (Eq 3), and (iii) clinically meaningful difference (CMD) (Eq 4) [29]:

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

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

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

The TEM% scores were qualitatively interpreted as: excellent: <5%; good: 5–10%; poor: >10%. The CMD was used to evaluate the smallest detectable difference that would need to be overcome to conclude that a true change has occurred. A generalised linear model was used to evaluate mean differences between groups (2 levels: Con vs. PFP), and devices (2 levels: ID vs. SG) where each participant was treated as a random effect in the model. Post-hoc analyses entailed the use of paired sample t-tests with a Holm correction to adjust for multiple comparisons. Standardised mean differences were calculated as Hedges' *g*, the magnitude of which was qualitatively interpreted as: trivial: <0.2; small: 0.2–0.6; moderate: 0.6–1.2; and large: > 1.2 [30].

The concurrent validity was determined by using linear regression where the Pearson correlation coefficient (*r*), coefficient of determination (*r*²), standard error of the estimate (SEE), and the slope of the regression lines were evaluated. The magnitude of the correlation coefficients were qualitatively interpreted as follows: negligible: 0.00–0.10; weak: 0.10–0.39; moderate: 0.40–0.69; strong: 0.70–0.89; and very strong: 0.90–1.00 [31]. Bland-Altman analyses were used to determine the systematic bias between the SG (reference measure) and the ID (criterion measure) [32]. For both the regression and Bland-Altman analyses, the point estimates were evaluated for potential outliers using Cook's distance where potential outliers (PO) were flagged when the Cook's distance exceeded a given threshold calculated as: 4/*n* (where *n* is the number of observations) [33]. All statistical analyses were completed using R [34].

Results

The results pertaining to the relative (ICC_{3,1}) and absolute (TEM) reliability as well as CMD of the SG are highlighted in Fig 2. Generally, the SG shows excellent relative and absolute reliability both at baseline and following the 8-week training interval. The ICC point estimates exhibit fairly narrow confidence intervals for the control group, but marginally longer intervals for the PFP group, indicating greater variability in torque measures. In all instances the CMD exceeded the TEM which implies that the SG had sufficient sensitivity to measure a true change in isometric quadricep strength (i.e. high signal-to-noise ratio).

Simple interaction effects for both between-group and within-group differences in mean torque values are shown in Fig 3. Mean differences between devices and group are shown in Fig 2A, whereas the standardised effect size with 95% CI and corresponding uncertainty density distribution for the point estimates are shown in Fig 2B.

For the concurrent validity, strong to very strong correlations were evident between the SG and ID torque measurements (*r* = 0.88–0.90, SEE = 0.05–0.07 Nm) (see Fig 4). Based on the slope analysis however, it is important to note that the torque values from the SG are typically lower than that obtained from the ID (slope = 0.65–0.77).

The results from the Bland-Altman analysis confirm the measurement bias between the SG and the ID whereby an average systematic bias of 16.3–28.8 Nm is evident in favour of the ID (see Fig 5).

Within-group differences in peak torque for each group and each device are shown in Fig 6. Although no statistically significant differences are evident, it is important to emphasise the

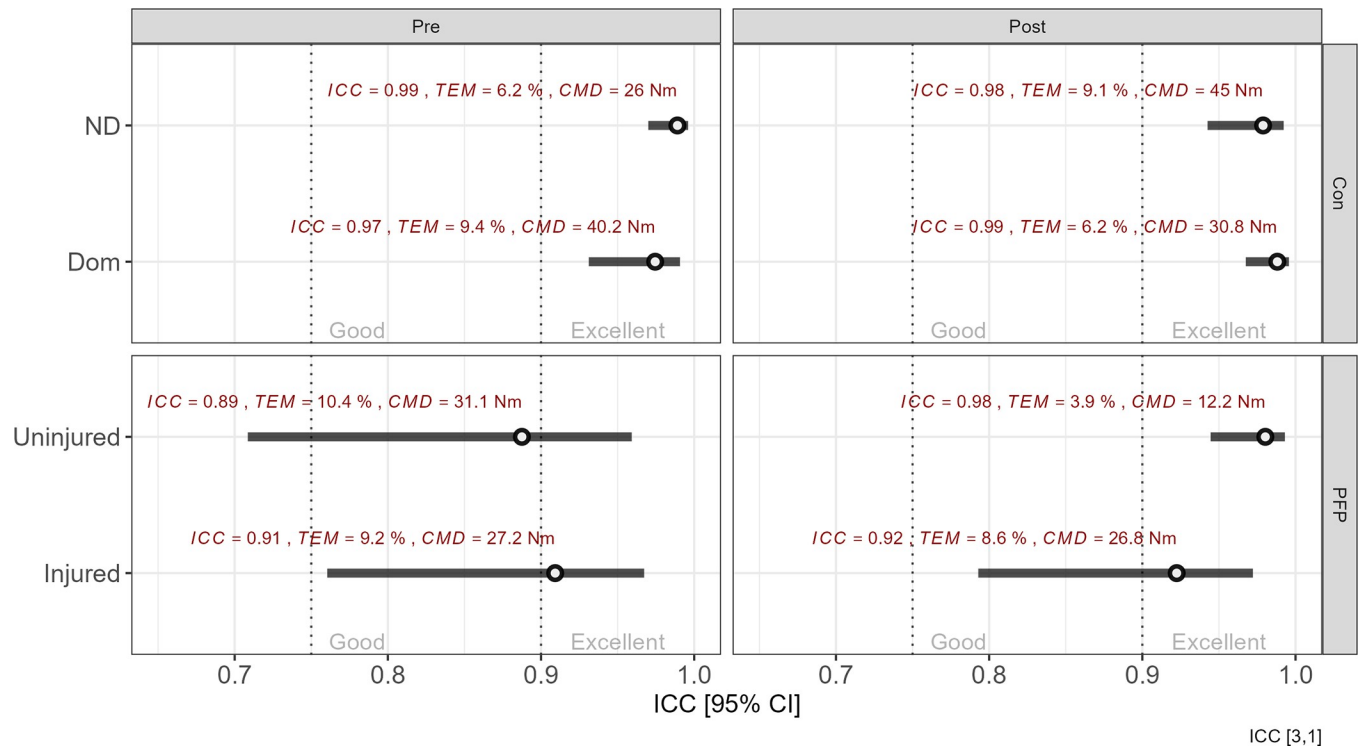


Fig 2. Between-group and within-group intra-class correlation coefficients. The ICC values are shown with their 95% confidence intervals. Above each point estimate are shown the precise ICC values, typical error of measurement (TEM%) and clinically meaningful difference (CMD) to provide greater context for the reliability score. Vertical dotted lines represent thresholds for excellent (ICC > 0.90) and good (ICC: 0.70–0.90). Note: ND = non-dominant; Dom = dominant; Con = control group; PFP = patellofemoral pain group.

<https://doi.org/10.1371/journal.pone.0298570.g002>

inter-individual variability highlighted by the colour-gradients which indicate the magnitude of participant-specific torque improvements in quadriceps strength between time points.

Discussion

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

Reliability

The SG showed excellent relative and absolute reliability in both the control and PFP groups (ICC > 0.90 and TEM < 10%) in all instances apart from the uninjured limb of the PFP group. The large ICC values obtained in the present study compared favourably with those from a separate study where belt-stabilisation was used with an HHD (ICC = 0.62 to 0.96) [35] as well as study focusing on the development of a novel affordable strain gauge (ICC = 0.99) [20]. Similarly, the TEM% scores of the present study were substantially better those reported by Martins et al. [35] who recorded TEM% of 12% for a belt-stabilised HHD and Chamorro et al. [18]

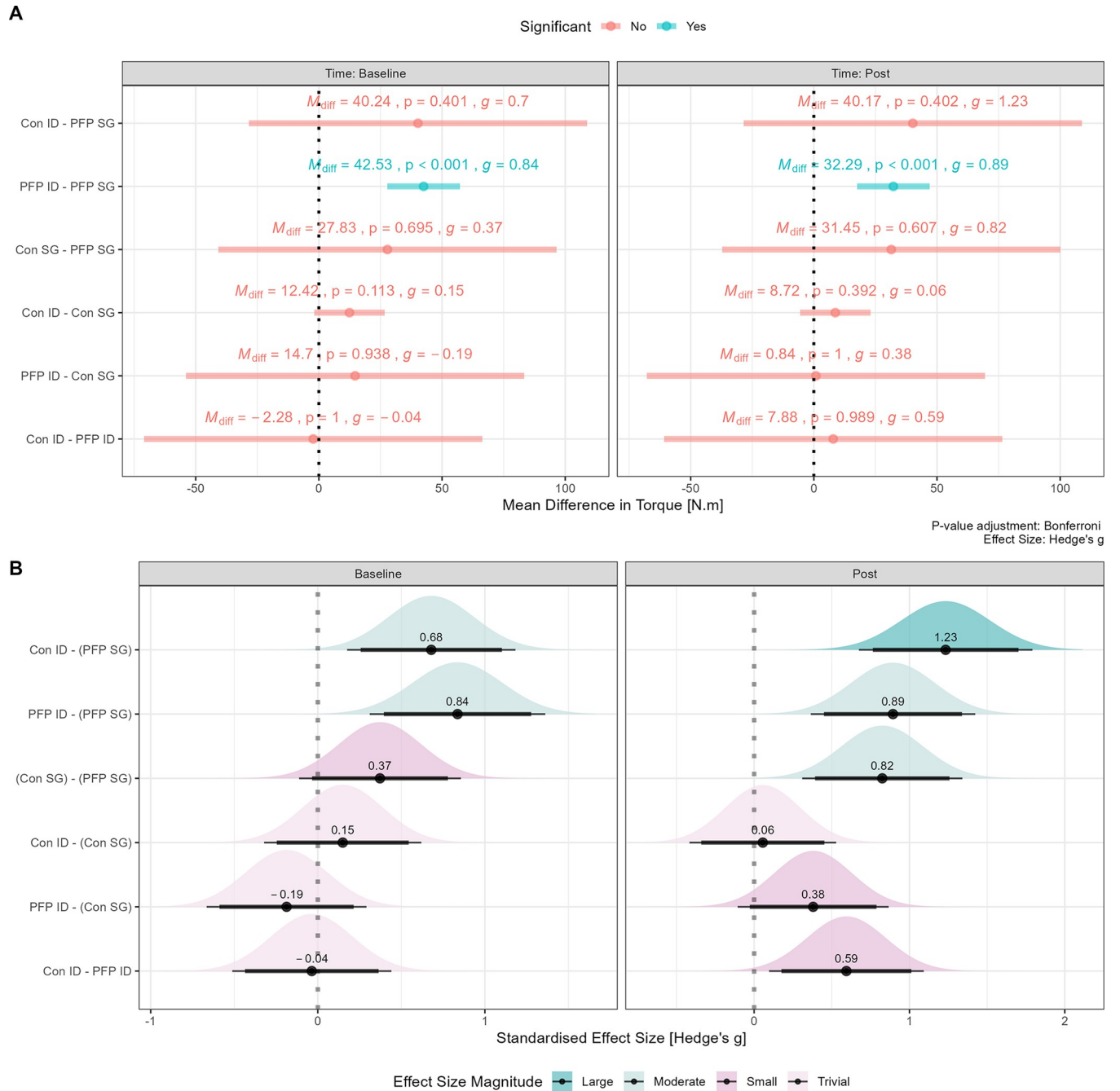


Fig 3. Absolute and relative within-group and between-group mean differences. Panel A: Mean differences with 95% CI; panel B: standardised mean effect size (Hedge's g) with 90% CI (thick black line) and 95% CI (thin black line) as well as density estimates to highlight the uncertainty in the point estimate. Note: M_{diff} = mean difference; g = Hedges' g effect size; Con = control group; PFP = patellofemoral pain group; ID = isokinetic dynamometer; SG = strain gauge.

<https://doi.org/10.1371/journal.pone.0298570.g003>

who yielded values ranging between 4–15% for measurements without stabilisation, both for knee extension strength. The ICC point estimates showed narrow confidence intervals in which the upper and lower limits fell within the margin defined as *excellent* in most instances. However, longer confidence intervals were observed in the PFP group during the initial set of testing indicated that the PFP group had greater variability in torque measures in the injured as well as the uninjured limbs compared to the control group. Torque variability for knee

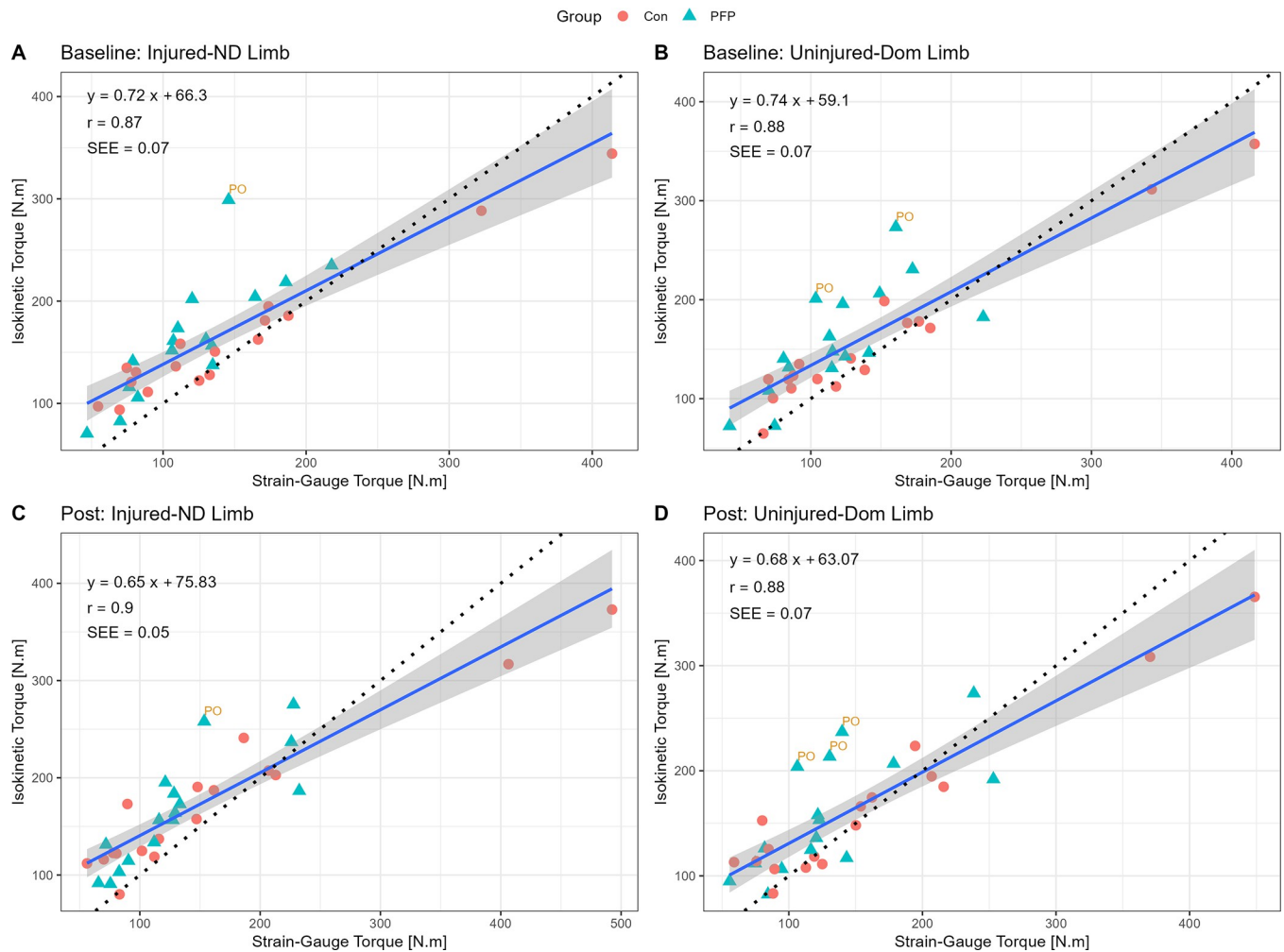


Fig 4. Linear regression between load-cell and Cybex for isometric knee extension torque. PO = potential outlier; SEE = standard error of the estimate; ND = non-dominant; Dom = dominant.

<https://doi.org/10.1371/journal.pone.0298570.g004>

extensors is present in other knee injuries as well [36], and in this case the variability may be accounted for in the PFP group on the basis that PFP patients tend to have impaired quadriceps function often ascribed to impaired vastus medialis oblique firing [37, 38] and reduced eccentric control [39]. However, the variability improved notably in the uninjured limb following the intervention period, possibly due to enhanced motor learning which tends to occur in injured populations [40].

It should be noted that the reliability of the SG, as with any device, is dependent on the set-up and therefore clinicians should be vigilant in following the set-up instructions set out in the procedures section and keep the set-up consistent when doing testing and re-testing in clinical settings, to ensure reliable results. Belt-stabilisation seems to improve reliability in SG-based devices and is recommended for use by practitioners [26, 35].

Clinically meaningful difference

The CMD within the present study ranged between 12–45 Nm, implying that the differences in torque production by the quadriceps musculature between limbs or over the course of

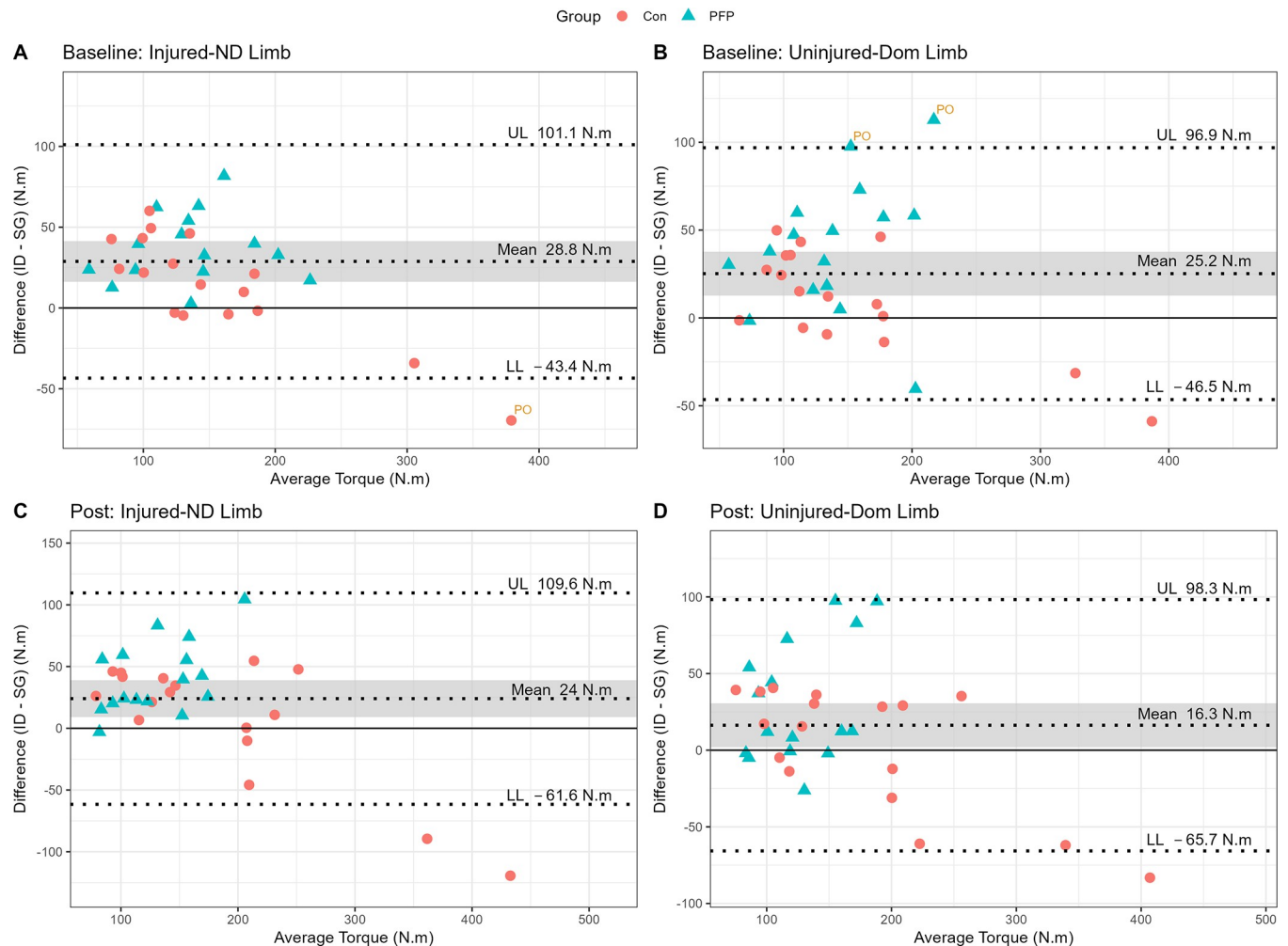


Fig 5. Bland-Altman plots for the bias (95% CI) between instruments and the limits of agreement. Note: UL = upper limit; LL = lower limit; ND = non-dominant; Dom = dominant; PO = potential outlier; ID = isokinetic dynamometer; SG = strain gauge.

<https://doi.org/10.1371/journal.pone.0298570.g005>

rehabilitation can be detected with the SG. Previous research has shown that an HHD exhibited comparable, although marginally lower CMD values of 17–27 Nm [41], with the ID yielding the lowest and most consistent CMD of approximately 27 Nm [42].

Group differences in torque production

An intriguing finding of the present study relates to the simple interaction effects regarding within-group (e.g., ID vs. SG) and between-group (e.g., control vs. PFP) differences for quadriceps torque (see Fig 3A and 3B). The mean differences in torque production between devices and groups are classified as trivial-to-moderate, and typically range -2.28 Nm to 27.83 Nm ($p = 0.113$ – 0.999). In most instances the point estimates of the mean difference were not significantly different from zero largely due to the longer confidence intervals. Such a result generally implies, at least in principle, similarities in torque production between groups and between devices. The standardised effect sizes (together with the 95% CI) show that, despite a lack of statistical significance, the magnitude of the mean differences are likely to be meaningful, especially for the PFP group which consistently yielded higher torque values on the ID

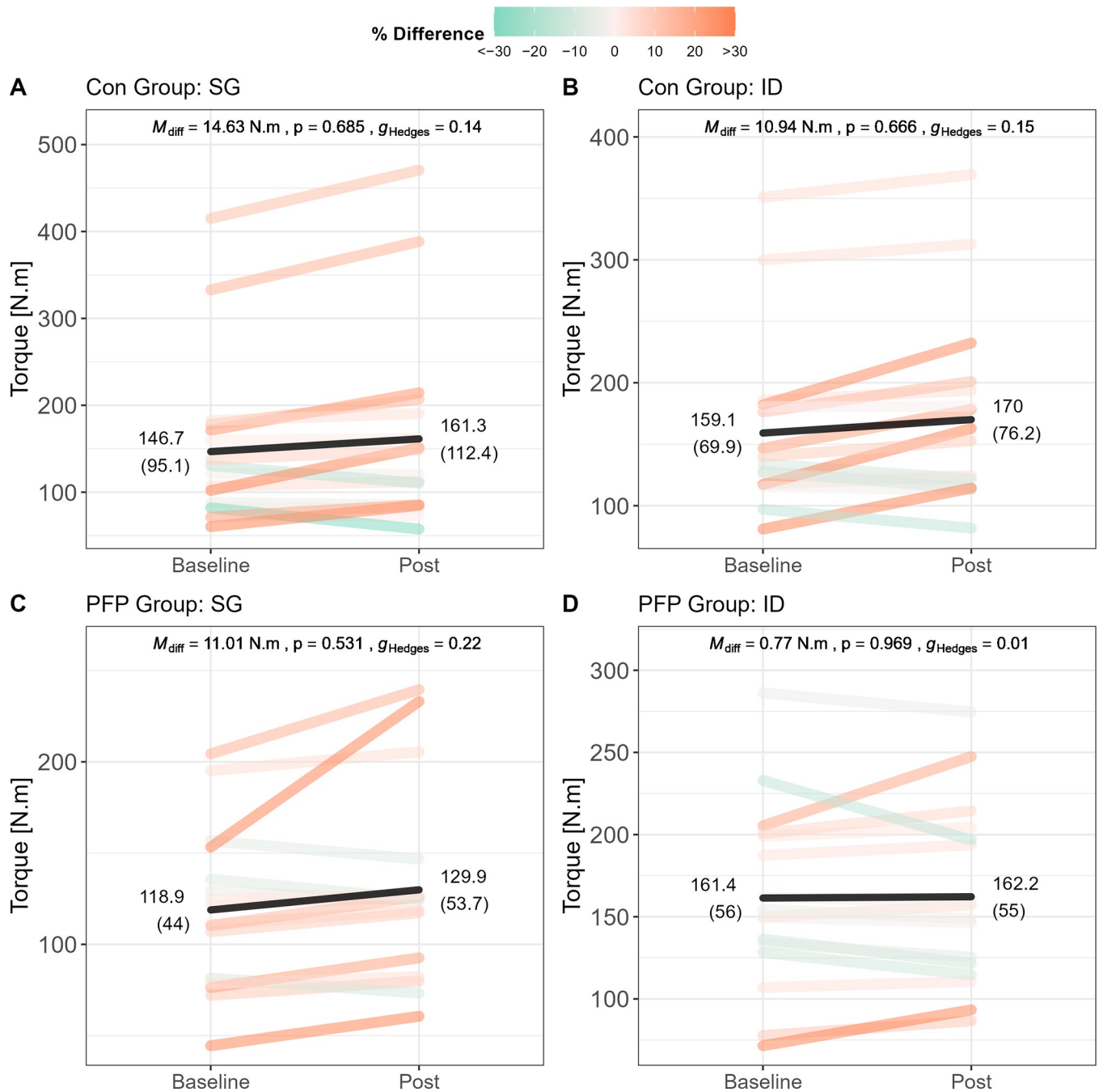


Fig 6. Within-group differences in peak torque. Pre-post differences are shown for the control group as evaluated by the SG (panel A) and ID (panel B). Pre-post differences are shown for the PFP group as evaluated by the SG (panel C) and ID (panel D). Line segments are coloured based on the percentage difference (Post-Pre/Pre \times 100) where improvements are shifted towards orange, and decrements are shifted towards green. The magnitude of the difference is captured by the intensity of the colour gradient (larger differences are darker, smaller differences are lighter). Black lines indicate the group mean for a given time point. Note: M_{diff} = mean difference; g_{Hedges} = Hedges' g ; Con = control group; PFP = patellofemoral pain group; SG = strain gauge.

<https://doi.org/10.1371/journal.pone.0298570.g006>

compared to the SG ($M_{diff} = 43.53$ Nm, $p < 0.001$). The interaction effects should be of interest to practitioners on the basis that the interchangeability of device measurements must be considered when interpreting results and making clinical inferences [42]. Given that, at least in some instances, practitioners might receive isokinetic results for a specific PFP patient or a

referring clinician, these results should be verified when using the SG in the clinical setting so that more precise interpretations can be made when evaluating temporal changes in quadriceps strength. Moreover, the same device should be used by the same clinician to ensure adequate consistency in readings and insights of whether true changes have accrued.

Validity

The results of the preset study suggest that the concurrent validity of the SG, when compared to the ID, appears to be strong to very strong, although a few discrepancies are noteworthy. The torque values from the SG are typically lower than those from the ID, with the exception of torque values exceeding ~300 Nm (see Figs 4 and 5). It is therefore reasonable to state that the SG might be of better use in injured and non-athletic populations that will most likely produce lower torque values compared to healthy, uninjured, or very athletic participants that might produce substantially higher torque values. The validity of the SG was greater than that of an HHD which yielded poorer scores even when stabilised with a belt, with correlations ranging between $r = 0.3$ – 0.8 [18, 43]. The SG is not a perfectly valid tool given the bias in torque readings and moderate-to-large limits of agreement (LoA: 72–86 Nm) and is therefore unlikely to replace the ID for absolute values. More specifically, there is evidence of a systematic bias in favour of the ID (see Fig 5), confirming that the SG typically measures lower torques than the ID on average. These results correspond with Martins et al. [35], who showed that knee extension tested with a belt-stabilised HHD, exhibited similar mean difference in torque production, with bias towards the ID. The HHD also shows, on average, wide LoA values (33.59%, CI_{95%} [23.91%, 43.26%]) for knee extension in other literature [18], which was echoed by Oranchuk et al. [44] who also showed fairly wide LoA for their experimental strain gauge (LoA: 150–200 Nm) during isometric quadriceps testing in a group of healthy participants. The wide LoA therefore support the notion that the SG is not valid enough to replace values obtained from the ID but given the excellent reliability, coupled with affordability, would still be an exceptional tool for the evaluation of isometric quadriceps strength. It is also important to highlight that isokinetic norms should not be used to make clinical interpretations when comparing these to values derived from the SG. It would be important to develop independently generated normative data for SG-derived torque values to facilitate decision-making across different joints and population groups.

Temporal changes in torque

Finally, although the mean peak torque values did not change significantly from the baseline to follow-up testing ($M_{diff} = 0.77$ – 14.63 Nm, $p = 0.531$ – 0.969), it should be noted that there were substantial individual improvements both within- and between groups (see Fig 6). The variability in individual responses underscores the potential inadequacy of the intervention programmes followed by the groups which were beyond the control of the current study. Whether more focused and targeted interventions or prolonged rehabilitation timelines (> 8 weeks) would elicit more favourable outcomes would require further research, especially in those with PFP. It should however be noted that there were consistencies in the measured values between devices across time, implying that the SG indeed has exceptional utility as a measurement tool for evaluating isometric quadriceps strength.

Limitations

Given the strengths of the present study, it is also important to underscore some limitations. Firstly, participants could follow any rehabilitation program at any practitioner of their choice, and therefore could not control the details associated with targeted quadriceps strengthening

as this was beyond the scope of the present study. Secondly, the ratios of males versus females in the groups differed, although consistency within groups was more important for this study than consistency between groups. Thirdly, participants were not divided into different groups according to the magnitude of quadriceps strength deficits. This latter point could provide more nuanced information on strength improvements over time in future studies.

Conclusions

The objectives of this study were to determine the validity and reliability of a low-cost SG for use in a clinical setting for measuring knee extension strength in those with PFP. The SG exhibited excellent reliability, and was deemed sensitive enough to detect clinically meaningful differences over time in both healthy individuals as well as those with PFP. The reliability of the SG is dependent on the set-up of the individual, and therefore practitioners should take care to complete the set-up as described and ensure consistency across every testing session. The SG lacks validity and is therefore unlikely to be an adequate surrogate for isokinetic dynamometry. However, given the trade-off in (i) costing associated with the ID and the SG, (ii) the importance of evaluating and tracking changes in knee extension strength in those with PFP, and (iii) the excellent reliability of the SG, the utility of the SG as a viable assessment tool is advocated, especially in resource-restricted settings.

Author Contributions

Conceptualization: Germari Deysel, Mariaan van Aswegen, Mark Kramer.

Data curation: Germari Deysel, Mark Kramer.

Formal analysis: Mark Kramer.

Methodology: Germari Deysel.

Project administration: Germari Deysel.

Supervision: Mark Kramer.

Visualization: Mark Kramer.

Writing – original draft: Germari Deysel, Mariaan van Aswegen, Mark Kramer.

Writing – review & editing: Germari Deysel, Mariaan van Aswegen, Mark Kramer.

References

1. Crossley KM, Stefanik JJ, Selfe J, Collins NJ, Davis IS, Powers CM, et al. 2016 Patellofemoral pain consensus statement from the 4th International Patellofemoral Pain Research Retreat, Manchester. Part 1: Terminology, definitions, clinical examination, natural history, patellofemoral osteoarthritis and patient-reported outcome measures. *Br J Sports Med* [Internet]. 2016; 50(14):839–43. Available from: <https://bjsm.bmj.com/content/50/14/831> <https://doi.org/10.1136/bjsports-2016-096384> PMID: 27343241
2. Hall R, Barber Foss K, Hewett TE, Myer GD. Sport specialization's association with an increased risk of developing anterior knee pain in adolescent female athletes. *J Sport Rehabil*. 2015; 24(1):31–5. <https://doi.org/10.1123/jsr.2013-0101> PMID: 24622506
3. Dey P, Callaghan M, Cook N, Sephton R, Sutton C, Hough E, et al. A questionnaire to identify patellofemoral pain in the community: An exploration of measurement properties. *BMC Musculoskeletal Disorders*. 2016; 17(1):e237. <https://doi.org/10.1186/s12891-016-1097-5> PMID: 27245443
4. Roush JR, Bay RC. Prevalence of anterior knee pain in 18–35 year-old females. *J Sports Phys Ther*. 2012; 7(4):396–401. PMID: 22893859
5. Kunene SH, Ramklass S, Taukobong NP. Anterior knee pain and its intrinsic risk factors among runners in under-resourced communities in Ekurhuleni, Gauteng. *S Afr J Physiother*. 2018; 74(1):e452. <https://doi.org/10.4102/sajp.v74i1.452> PMID: 30349876

6. Coburn SL, Barton CJ, Filbay SR, Hart HF, Rathleff MS, Crossley KM. Quality of life in individuals with patellofemoral pain: A systematic review including meta-analysis. *Physical Therapy in Sport* [Internet]. 2018; 33:96–108. Available from: <https://doi.org/10.1016/j.ptsp.2018.06.006> PMID: 30059951
7. Collins NJ, Bierma-Zeinstra SMA, Crossley KM, Van Linschoten RL, Vicenzino B, Van Middelkoop M. Prognostic factors for patellofemoral pain: a multicentre observational analysis. *Br J Sports Med* [Internet]. 2013; 47(4):227–33. Available from: <https://pubmed.ncbi.nlm.nih.gov/23242955/> <https://doi.org/10.1136/bjsports-2012-091696> PMID: 23242955
8. Reijnders L, Van de Groes SAW. The quality of life of patients with patellofemoral pain—a systematic review. *Acta Orthop Belg*. 2020; 86(4):678–87. PMID: 33861916
9. Tan SS, van Linschoten RL, van Middelkoop M, Koes BW, Bierma-Zeinstra SM, Koopmanschap MA, et al. Cost-utility of exercise therapy in adolescents and young adults suffering from the patellofemoral pain syndrome. *Scand J Med Sci Sports*. 2010; 20(4):568–79. <https://doi.org/10.1111/j.1600-0838.2009.00980.x> PMID: 19706002
10. Nimon G, Murray D, Sandow M, Goodfellow J. Natural History of Anterior Knee Pain: A 14- to 20-Year Follow-up of Nonoperative Management [Abstract]. *Journal of Pediatric Orthopaedics* [Internet]. 1998 [cited 2023 Apr 17]; 18(1):118–22. Available from: https://journals.lww.com/pedorthopaedics/Abstract/1998/01000/Natural_History_of_Anterior_Knee_Pain_A_14_to_21.aspx
11. Thomas MJ, Wood L, Selfe J, Peat G. Anterior knee pain in younger adults as a precursor to subsequent patellofemoral osteoarthritis: a systematic review. *BMC Musculoskelet Disord* [Internet]. 2010 [cited 2023 Apr 17]; 11(1):e201. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2944218/> <https://doi.org/10.1186/1471-2474-11-201> PMID: 20828401
12. Van Middelkoop M, Bennell KL, Callaghan MJ, Collins NJ, Conaghan PG, Crossley KM, et al. International patellofemoral osteoarthritis consortium: Consensus statement on the diagnosis, burden, outcome measures, prognosis, risk factors and treatment. *Semin Arthritis Rheum*. 2018; 47(5):666–75. <https://doi.org/10.1016/j.semarthrit.2017.09.009> PMID: 29056348
13. Boling MC, Padua DA, Marshall SW, Guskiewicz K, Pyne S, Beutler A. A Prospective Investigation of Biomechanical Risk Factors for Patellofemoral Pain Syndrome The Joint Undertaking to Monitor and Prevent ACL Injury (JUMP-ACL) Cohort. *Am J Sports Med* [Internet]. 2009; 37(11):2108–16. Available from: <http://ajs.sagepub.com/supplemental/> <https://doi.org/10.1177/0363546509337934> PMID: 19797162
14. Amin S, Baker K, Niu J, Clancy M, Goggins J, Guermazi A, et al. Quadriceps Strength and the Risk of Cartilage Loss and Symptom Progression in Knee Osteoarthritis. *Arthritis Rheum* [Internet]. 2009 [cited 2023 Apr 17]; 60(1):189–98. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653642/> <https://doi.org/10.1002/art.24182> PMID: 19116936
15. Willy RW, Hoglund LT, Barton CJ, Bolgla LA, Scalzitti DA, Logerstedt DS, et al. Patellofemoral pain. *J Orthop Sports Phys Ther*. 2019; 49(9):CPG1–95. <https://doi.org/10.2519/jospt.2019.0302> PMID: 31475628
16. Janik F, Toulotte C, Seichepine AL, Masquelier B, Barbier F, Fabre C. Isometric Strength Database for Muscle Maximal Voluntary Endurance Field Tests: Normative Data. *Sports Med* [Internet]. 2021; 7(1):e47. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8273050/> <https://doi.org/10.1186/s40798-021-00338-2> PMID: 34250556
17. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held Dynamometry Correlation With the Gold Standard Isokinetic Dynamometry: A Systematic Review. *PM&R* [Internet]. 2011; 3(5):472–9. Available from: <https://onlinelibrary-wiley-com.nwulib.nwu.ac.za/doi/full/10.1016/j.pmrj.2010.10.025> PMID: 21570036
18. Chamorro C, Armijo-Olivo S, De La Fuente C, Fuentes J, Javier Chiroso L. Absolute reliability and concurrent validity of hand held dynamometry and isokinetic dynamometry in the hip, knee and ankle joint: Systematic review and meta-analysis. *Open Med* [Internet]. 2017; 12(1):359–75. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5651404/> <https://doi.org/10.1515/med-2017-0052> PMID: 29071305
19. Deones VL, Wiley SC, Worrell T. Assessment of quadriceps muscle performance by a hand-held dynamometer and an isokinetic dynamometer. *JOSPT* [Internet]. 1994; 20(6):296–301. Available from: <https://pubmed.ncbi.nlm.nih.gov/7849749/> <https://doi.org/10.2519/jospt.1994.20.6.296> PMID: 7849749
20. Juneau C, Oranchuk DJ, Cahill M, Forster JW, Diewald S. Reliability and Utility of Load-Cell Derived Force–Time Variables Collected During a Constrained and Unconstrained Isometric Knee Extension Task on a Plinth. *Journal of Science in Sport and Exercise*. 2024; 6:81–9.
21. Christensen LB, Johnson RB, Turner LA. Research methods, design and analysis. 12th ed. Vol. 18. Harlow: Pearson Education Limited; 2015.
22. Faul F, Erdfelder E, Lang A, Buchner A. Statistics textbooks in the social. *Behavior Research Methods* [Internet]. 2007; 39(2):175–91. Available from: <https://link.springer.com/content/pdf/10.3758/BF03193146.pdf>

23. Howe TE, Dawson LJ, Syme G, Duncan L, Reid J. Systematic review Evaluation of outcome measures for use in clinical practice for adults with musculoskeletal conditions of the knee: A systematic review. *Man Ther.* 2012; 17(2):100–18.
24. Watson CJ, Propps M, Ratner J, Zeigler DL, Horton P, Smith SS. Reliability and responsiveness of the lower extremity functional scale and the anterior knee pain scale in patients with anterior knee pain. *J Orthop Sports Phys Ther* [Internet]. 2005; 35(3):136–46. Available from: <https://pubmed.ncbi.nlm.nih.gov/15839307/> <https://doi.org/10.2519/jospt.2005.35.3.136> PMID: 15839307
25. Hasan S. Quadriceps Femoris Strength Training: effect of Neuromuscular Electrical Stimulation Vs Isometric Exercise in Osteoarthritis of Knee. *Indian J Physiother Occup Therapy.* 2015; 9(3):129–34.
26. McNamara S, Hevia EG, St. Louis R, Cho W, Lee S, Moyne M, et al. Isometric quadriceps strength test device to improve the reliability of handheld dynamometry in patient with anterior cruciate ligament injury. In Minneapolis, MN, USA: ResearchGate; 2018. Available from: <http://asmedigitalcollection.asme.org/BIOMED/proceedings-pdf/DMD2018/40789/V001T09A001/2788227/v001t09a001-dmd2018-6848.pdf>
27. Aby F. Serotonergic neurons of the nucleus raphe Magnus in the control of nociceptive transmission in the dorsal horn of the spinal cord: an optogenetic study in different pathophysiological contexts. [Nouvelle-Aquitaine, France]: University of Bordeaux; 2019.
28. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* [Internet]. 2016; 15(2):155–155. Available from: /pmc/articles/PMC4913118/ <https://doi.org/10.1016/j.jcm.2016.02.012> PMID: 27330520
29. Schwenk M, Gogulla S, Englert S, Czempik A, Hauer K. Test–retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. *Physiol Meas* [Internet]. 2012 Oct [cited 2024 Jan 8]; 33(11):1931. Available from: <https://doi.org/10.1088/0967-3334/33/11/1931> PMID: 23110800
30. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc.* 2009 Jan; 41(1):3–13. <https://doi.org/10.1249/MSS.0b013e31818cb278> PMID: 19092709
31. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia* [Internet]. 2018 May [cited 2024 Jan 8]; 126(5):1763. Available from: https://journals.lww.com/anesthesia-analgesia/fulltext/2018/05000/correlation_coefficients_appropriate_use_and.50.aspx <https://doi.org/10.1213/ANE.0000000000002864> PMID: 29481436
32. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)* [Internet]. 2015 Jun 5 [cited 2023 Dec 19]; 25(2):141–51. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4470095/> <https://doi.org/10.11613/BM.2015.015> PMID: 26110027
33. Hebbali A. *rsquaredacademy/olsrr* [Internet]. Rsquared Academy; 2023 [cited 2024 Jan 17]. Available from: <https://github.com/rsquaredacademy/olsrr>
34. R Core Team. R: The R Project for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; [cited 2024 Jan 8]. Available from: <https://www.r-project.org/>
35. Martins J, Rodrigues Da Silva J, Rodrigues Barbosa Da Silva; D ´ Eborá Bevilaqua-Grossi M. Reliability and Validity of the Belt-Stabilized Handheld Dynamometer in Hip-and Knee-Strength Tests. *J Athl Train* [Internet]. 2017; 52(9):809–19. Available from: www.natajournals.org <https://doi.org/10.4085/1062-6050-52.6.04> PMID: 28787180
36. Goetschius J, Hart JM. Knee-Extension Torque Variability and Subjective Knee Function in Patients with a History of Anterior Cruciate Ligament Reconstruction. *J Athl Train.* 2016 Jan; 51(1):22–7. <https://doi.org/10.4085/1062-6050-51.1.12> PMID: 26720104
37. Van Tiggelen D, Cowan S, Coorevits P, Duvigneaud N, Witvrouw E. Delayed vastus medialis obliquus to vastus lateralis onset timing contributes to the development of patellofemoral pain in previously healthy men: a prospective study [Abstract]. *Am J Sports Med.* 2009; 37(6):1099–105.
38. Pal S, Draper CE, Fredericson M, Gold GE, Delp SL, Beaupre GS, et al. Patellar maltracking correlates with vastus medialis activation delay in patellofemoral pain patients. *Am J Sports Med.* 2011; 39(3):590–8. <https://doi.org/10.1177/0363546510384233> PMID: 21076015
39. Eapen C, Nayak CD, Pazhyaottiyil Zulfequer C. Effect of Eccentric Isotonic Quadriceps Muscle Exercises on Patellofemoral Pain Syndrome: An Exploratory Pilot Study. *Asian J Sports Med* [Internet]. 2011 [cited 2023 Jun 3]; 2(4):227–34. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3289225/> <https://doi.org/10.5812/asjasm.34747> PMID: 22375243
40. Rodrigues-da-Silva JM, de Rezende MU, Spada TC, da Silva Francisco L, Greve JMD, Ciolac EG. Effects of Motor Learning on Clinical Isokinetic Test Performance in Knee Osteoarthritis Patients. *Clinics (Sao Paulo)* [Internet]. 2017 [cited 2024 Jan 12]; 72(4):202–6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5401619/> [https://doi.org/10.6061/clinics/2017\(04\)02](https://doi.org/10.6061/clinics/2017(04)02) PMID: 28492718

41. Knols RH, Aufdemkampe G, de Bruin ED, Uebelhart D, Aaronson NK. Hand-held dynamometry in patients with haematological malignancies: Measurement error in the clinical assessment of knee extension strength. *BMC Musculoskelet Disord* [Internet]. 2009 Mar 9 [cited 2024 Jan 12]; 10:31. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2662793/> <https://doi.org/10.1186/1471-2474-10-31> PMID: 19272149
42. de Araujo Ribeiro Alvares JB, Rodrigues R, de Azevedo Franke R, da Silva BGC, Pinto RS, Vaz MA, et al. Inter-machine reliability of the Biodex and Cybex isokinetic dynamometers for knee flexor/extensor isometric, concentric and eccentric tests. *Physical Therapy in Sport* [Internet]. 2015 Feb 1 [cited 2024 Jan 17]; 16(1):59–65. Available from: <https://www.sciencedirect.com/science/article/pii/S1466853X14000339> <https://doi.org/10.1016/j.ptsp.2014.04.004> PMID: 24913915
43. Hirano M, Katoh M, Gomi M, Arai S. Validity and reliability of isometric knee extension muscle strength measurements using a belt-stabilized hand-held dynamometer: a comparison with the measurement using an isokinetic dynamometer in a sitting posture. *Journal of Physical Therapy Science*. 2020; 32(2):120–4. <https://doi.org/10.1589/jpts.32.120> PMID: 32158074
44. Oranchuk D, Juneau C, Diewald S, Neville J, Cronin J. The Validity of a Portable Strain-Gauge Apparatus Versus a Commercial Isokinetic Dynamometer for Evaluating Knee Extension Kinetics. *International Journal of Sports Physical Therapy* [Internet]. 2024 Mar 2 [cited 2024 Apr 30];19(3). Available from: <https://ijspt.scholasticahq.com/article/92742-the-validity-of-a-portable-strain-gauge-apparatus-versus-a-commercial-isokinetic-dynamometer-for-evaluating-knee-extension-kinetics> <https://doi.org/10.26603/001c.92742> PMID: 38439775