Reliability and Criterion Validity of Knee Frontal Plane Projection Angles Measured Using the Technique Application

Lee T Atkins[1], Tyler Van Bastian TJ[2], Sterling D Eckert[3], Timothy R Hofmann[4], Emily L Pierce[5], Jim Richards[6], James Selfe[7], Michael Callaghan[7]

1. Department of Physical Therapy, Angelo State University, San Angelo, Texas, USA

2. New Mexico Orthopaedics, Rio Rancho, NM, USA

3. Fyzical Therapy and Balance Center, San Angelo, TX, USA

4. Dignity Health, Phoenix, AZ, USA

5. Carrel Clinic, Dallas, TX, USA

6. Allied Health Research unit, University of Central Lancashire, UK

7. Department of Health Professions, Manchester Metropolitan University, UK

Contact Information for Corresponding Author:
Lee Atkins
Email: lee.atkins@angelo.edu
Phone: 325-486-6865

ABSTRACT

Context: Abnormal knee frontal plane projection angles (FPPA) during movement have been associated with patellofemoral pain (PFP). As such, clinicians are interested in valid and reliable instruments suitable for broad-based clinical use that allow them to objectively measure such variables. Therefore, the purpose of the current study was to examine the criterion validity and reliability of knee FPPA measures obtained by clinicians using a free tablet application called Technique.

Design: validity/reliability study

Methods: To examine validity, the same raters measured ten, two-dimensional criterion reference angles at the first testing session. To examine reliability, the knee FPPA of sixteen subjects was measured by 6 raters (3 physical therapists and 3 student physical therapists) on two separate occasions while performing a single-limb stepdown task. Validity was investigated by calculating the 95% limits of agreement (LA), mean absolute differences, and Bland-Altman plots. Reliability was examined by calculating intraclass correlation coefficients (ICC) and the standard error of measure (SEM).

Results: For validity, the mean absolute difference between rater and criterion reference angle measures ranged from 0.20 to 0.90 degrees. 95% of expected errors between rater and criterion reference angle measures were 2.04 degrees or less. For reliability, the ICC values for inter- and intrarater reliability were excellent ranging from 0.994 to 0.998 with SEM ranging from 0.44 to 0.84 degrees.

Conclusions: These findings indicate that knee FPPA measures obtained during a single-limb stepdown task using the Technique tablet application are valid and reliable, and suitable for clinical use.

INTRODUCTION

Patellofemoral pain (PFP) is a multifaceted condition with an estimated annual prevalence of 22.7%.[1] Abnormal frontal plane knee kinematics leading to increased patellofemoral joint stress[2] is commonly assumed to be a major contributing factor to PFP.[3] Compared to control subjects, individuals with PFP exhibit altered frontal plane knee kinematics during a single-limb stepdown task.[4,5] It has also been reported that the knee frontal plane projection angle (FPPA) is a predictor of PFP development[6] and is significantly correlated with PFP severity[7] and hip muscle weakness.[8] The knee FPPA represents the degrees of non-collinearity between thigh and leg segments that project onto the frontal plane. Additionally, improvements in frontal plane knee kinematics have been associated with PFP improvements.[9] As such, it has been recommended clinicians examine the knee FPPA[10,11] during functional tasks such as single-limb stepdown[12] for individuals diagnosed with PFP. Thus, clinicians require an objective, reliable, and valid means for quantifying frontal plane knee kinematics for patients diagnosed with PFP.

Clinically, movement kinematics have been assessed using subjective real-time analyses that do not allow clinicians to quantify movement kinematics. In contrast, traditional lab assessment of movement kinematics has required complex and expensive equipment not practical for many clinics. However, it has been suggested that in a clinical setting, the use of complex and expensive equipment is only necessary if the information needed is otherwise unattainable.[13] For example, researchers reported that hip kinematics recorded and measured using a commercially available digital video camera, smartphone running a free inclinometer application, and external monitor were both valid and reliable.[14] Technological advances have led to increased availability and use of tablet devices by clinicians. While using free applications, tablet devices can be used to record video and quantify kinematics, eliminating the need for a separate digital video camera and external monitor for displaying

the video. An example of a tablet application is Technique (Hudl Inc., Lincoln, NE). Once opened, users can record video within the Technique application. The recorded video can be replayed and specific frames of interest can be identified. Users can then use an angle drawing tool to measure angles (on the frame of interest) like the knee FPPA with nothing more than a tablet. Furthermore, for educational purposes, applications like Technique make it easy for clinicians to record and replay frame-by-frame 2D digital video and quantify kinematics on a particular frame of interest. Thus, considering the promise of this technology and to avoid misinterpreting clinical data, the validity and reliability of tablet applications such as Technique need to be established prior to broad-based clinical implementation.

Several studies have examined the reliability of measuring lower limb kinematics on 2D digital video recorded with a handheld device. For example, using an application called Kinesiocapture, King and Belyea reported intraclass correlation coefficients (ICC) for inter- and intrarater reliability for knee FPPA during landing as low as 0.45 and 0.44, respectively.[15] They also reported the standard error of measure between raters as high as 7.2 degrees, thus calling into question the clinical usefulness of this application.[15] Similarly, Neal et al. reported low to moderate reliability (ICC 0.31 – 0.71) for measuring lower limb kinematics during running using Technique.[16] It is possible in these studies that reliability may have improved if a few methodological changes were made. For example, the use of skin markers may improve repeatability of aligning the angle measurement tool[15], and the use of a stylus instead of a finger might improve precision in aligning the angle measurement tool.

Thus, it is plausible that incorporating a few simple changes such as skin markers and a stylus that the validity, reliability, and overall clinical usefulness of 2D kinematics measures made with a tablet application could be substantially improved over previous studies. Therefore, the purpose of the current study was to examine the criterion validity and reliability of the knee FPPA during a single-limb stepdown task measured using the Technique application, skin markers, and stylus. It was hypothesized

that the knee FPPA measured with Technique could be clinically suitable, but this needs to be initially explored using criterion validity and reliability.

METHODS

Participants

Subjects

A convenience sample of 16 subjects (8 males and 8 females) [mean (SD) age: 23.1 (2.1) years; height: 1.7 (0.1) m; mass: 71.4 (10.4) kg; body mass index: 23.6 (2.4) kg/m^2] were recruited through word-of-mouth from a university population. For inclusion, volunteers must have been at least 18 years of age, English-speaking, had a BMI less than 30 kg/m^2, and free of back or leg pain. Volunteers were excluded if they were pregnant, had an orthopaedic or neuromuscular condition that made it unsafe to perform a single-limb stepdown, or their healthcare provider had instructed them to avoid activities comparable to a single-limb stepdown.

Raters

A convenience sample of six clinician raters (three physical therapists and three student physical therapists) [mean (SD) age: 32.0 (13.2) years] were recruited through word-of-mouth from local orthopaedic physical therapy clinics and a university. The mean (SD) years practicing as physical therapists was 15.7 (17.8) years and the student physical therapists were in their first year of a physical therapy graduate program. For inclusion, volunteers must have been English-speaking and either a physical therapist with a license in good standing or a student physical therapist enrolled in an

accredited physical therapy program. Volunteers were excluded for uncorrected visual impairments or prior experience using the Technique to quantify movement.

The Institutional Review Board at Angelo State University approved this study, and all participants gave written informed consent.

Procedures

Subject Procedures

Subjects donned athletic shoes (Asics, Irvine, CA), black spandex shorts and sports top (females only) to reduce rater recall at subsequent testing sessions. Next, the limb of interest was determined by coin flip and reflective markers (14 mm diameter) attached midway between femoral epicondyles, midway between malleoli, and on the proximal thigh on a line connecting the anterior superior iliac spine and knee marker (Figure 1).[10]

A tablet (Apple iPad Pro, model A1670; Apple Inc., Cupertino CA) was then secured to a tripod, aligned perpendicular to the frontal plane (yaw) and leveled in two planes (pitch and roll)(Figure 1).[17] The tablet height was adjusted to the approximate height of the knee marker while the subject stood on the step (Figure 1). To minimize image distortion, the distance between subject and tablet was maximized and the image enlarged using the zoom feature so only the markers of interest, trunk, and the subjects' lower limbs were visible (Figure 1).[17] For the single-limb stepdown, subjects were instructed to lower themselves at self-selected speed until the heel of their non-stance limb touched the ground in front of the step, at which point they were to return to the starting position (Figure 1). Subjects were allowed practice to familiarize themselves with the task. Next, the Technique tablet application (v5.4.6.6) was used to record digital video (30 Hz) of each subject as they performed five

single-limb stepdown trials from a 20 cm step (Figure 1). The decision to standardize the step height was because for many daily tasks, the physical constraints of steps are not adjusted based on one's body height. For each subject a single trial was arbitrarily selected for a total of 16 trials to be assessed by the raters.

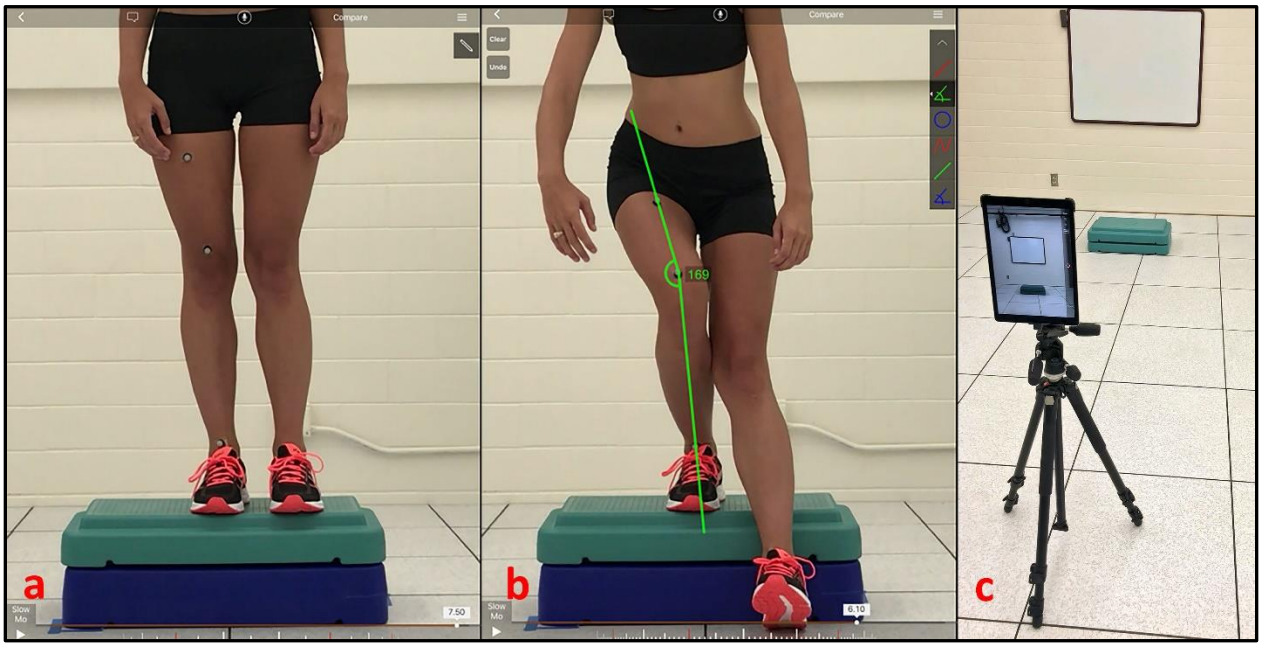

Figure 1. Still frame images of video capture (using the Technique application) demonstrating **a)** marker placement locations on subject, **b)** knee FPPA measurement by a rater using the angle drawing tool, and **c)** positioning of the iPad on an adjustable tripod relative to the step.

Criterion Validity

Criterion validity was examined by assessing the level of agreement between angles measured using Technique and those measured using an electrogoniometer (model SG150; Biometrics Ltd, Newport, UK). Each rater measured 10 criterion reference angles with Technique by using three non-

collinear markers placed on a planar surface oriented perpendicular to the tablet (Figures 2). The criterion reference measure of these angles was established by researchers using the electrogoniometer. Research indicates that the electrogoniometer provides accurate angle measures compared to an infrared motion capture system.[18]

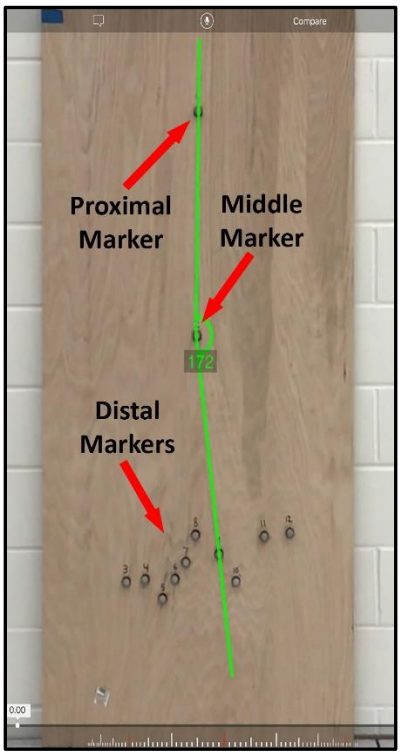

Figure 2. Screen capture within the Technique application of a single criterion reference angle (as measured by a rater) used to examine validity. Each rater assessed the 10 criterion reference angles by measuring degrees of non-collinearity between the line connecting the proximal and middle markers and the lines connecting the middle and each of the 10 distal markers (for a total of 10 criterion reference angles). For comparison to the rater measures, the same 10 angles were measured by researchers using an electrogoniometer.

Rater Procedures

At the first of two testing sessions, raters were provided instructions concerning the Technique application. Specifically, raters were instructed to open Technique on the iPad and select specific video frames of interest (within the application) for each subject. Next, they were instructed to use the zoom

function to enlarge the lower limb and align the angle measurement tool (a feature within Technique) with the reflective markers to quantify knee FPPA (Figure 3). Raters used a stylus (Apple Pencil 1st Generation; Apple Inc., Cupertino CA) to assist with aligning the angle measurement tool (Figure 3). Raters were allowed to practice measuring knee FPPA's on sample video frames. Next, raters measured the knee FPPA for each of the 16 subjects on the video frames of interest (when the subject's non-stance heel contacted the ground) as specified by a researcher. The knee FPPA measures from the first testing session were used to establish interrater reliability. Each rater assessed the knee FPPA using the same video frames at a second testing session to establish intrarater reliability. The two testing sessions were separated by at least seven days, and the video frame presentation order was randomized.

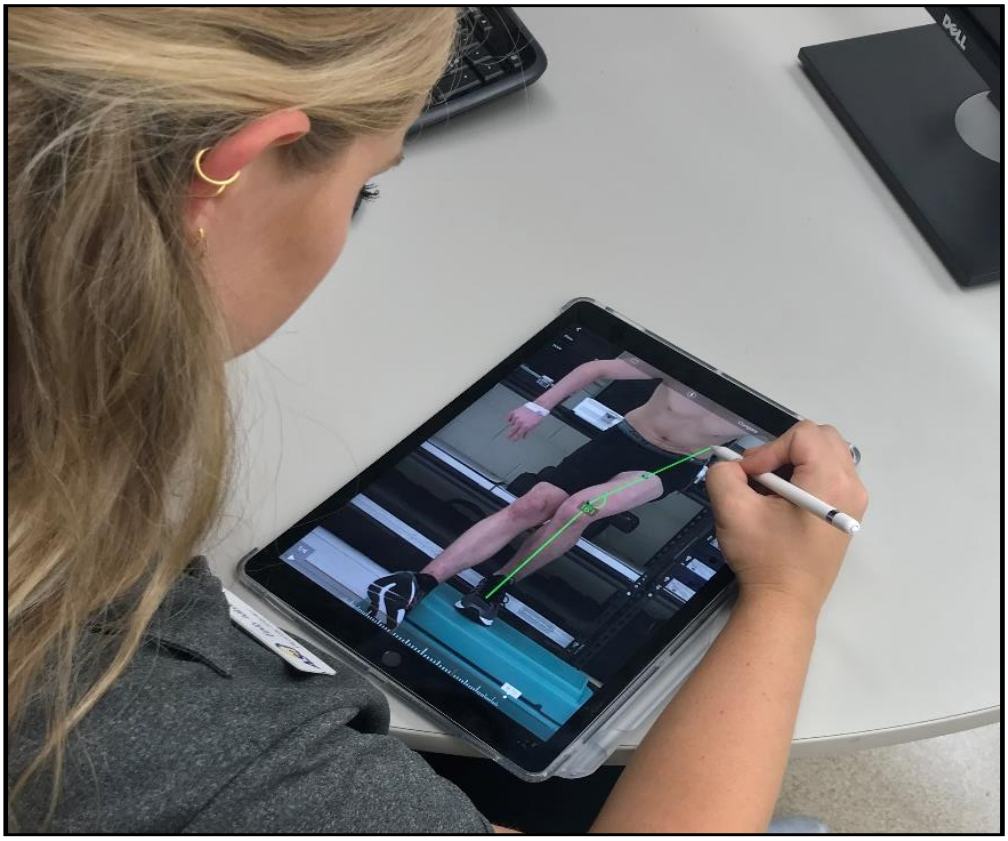

Figure 3. Rater using the angle measurement tool and a stylus to measure the knee frontal plane projection angle within the Technique application.

Statistical Analyses

Criterion Validity

Initially, Bland–Altman plots with 95% limits of agreement (LOA) were calculated using criterion reference angle measures obtained by raters via Technique and the researchers via the electrogoniometer.[19] Next, mean absolute difference (MAD) and mean difference (MD) between rater and reference angle values were calculated. Additionally, 95% limits of agreement (LA) were calculated to determine the range within which 95% of the differences between rater and reference measures would lie. The formula for calculating the 95% LA was LA = MD ± (2 × SD).[20] Based on research indicating 4.1° difference in knee FPPA between individuals with PFP and asymptomatic individuals[10] acceptable agreement was operationally defined as having 95% LOA that fell within a 2 degree range which is less than half of the aforementioned knee FPPA difference. All statistics were performed using SPSS (version 25; IBM Inc., Armonk NY) and Excel (2016, Microsoft Inc., Redmond WA).

Reliability

Intrarater and interrater reliability were examined by calculating intraclass correlation coefficients ($ICC_{(3,1)}$ and $ICC_{(2,1)}$, respectively), 95% confidence intervals, and standard error of measurement (SEM) [(SD × (1-ICC)]. For intrarater reliability, the range within which 95% of measurement errors were expected upon reassessment of knee FPPA was also estimated (± SEM × 1.96). Similarly, for interrater reliability, the range within which 95% of expected errors between raters was also estimated (± SEM × 1.96). Intraclass correlation coefficients were interpreted as excellent (greater than or equal to 0.90), good (between 0.75 and 0.90), and moderate to poor (less than or equal to 0.75).[20]

RESULTS

Criterion Validity

The mean (SD) rater and criterion reference angles are shown in Table 1. The mean criterion reference angle was -2.10 (15.82) degrees whereas the mean rater measures of the criterion reference angles ranged from -1.20 to -2.00 degrees (Table 1). The MAD between criterion and rater reference measures ranged from 0.20 to 0.90 degrees, while the 95% LA ranged from 1.04 to 2.04 degrees (Table 1). Additionally, based on the largest MD, 95% of differences between rater and criterion reference measures are expected to be 2.04 degrees or less (Table 1). Finally, visual inspection of Bland-Altman plots suggested no systematic errors were present in rater measures of criterion reference angles (Figure 4).

Table 1. Criterion Validity

	Criterion Reference	**Rater 1**	**Rater 2**	**Rater 3**	**Rater 4**	**Rater 5**	**Rater 6**
Mean ± SD[†] (°)	-2.10 ± 15.82	-1.20 ± 15.56	-2.00 ± 15.56	-1.80 ± 15.44	-1.90 ± 15.64	-1.90 ± 15.64	-1.90 ± 15.64
MAD (°)	---	0.90 ± 0.57	0.30 ± 0.48	0.30 ± 0.48	0.20 ± 0.42	0.20 ± 0.42	0.20 ± 0.42
95% LA (°)	---	2.04	1.27	1.27	1.04	1.04	1.04

[†]represents the mean (SD) of all 10 criterion reference angles.
SD, standard deviation; MAD, mean absolute difference, 95% LA, 95% limits of agreement.

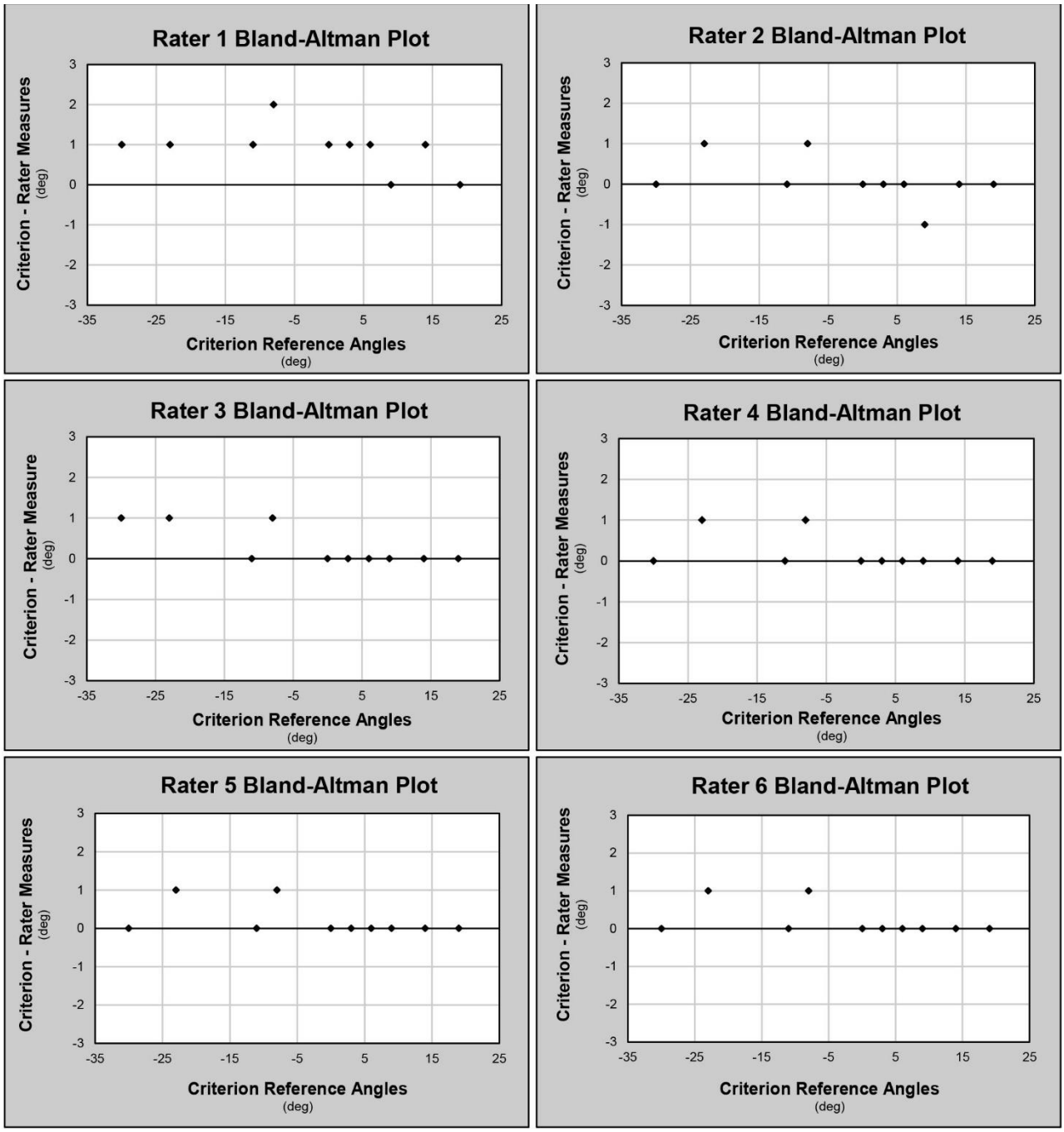

Figure 4. Bland-Altman plots for each rater. Positive y-axis values indicate the rater angle measure was less than the electrogoniometer criterion reference angle measure. Negative y-axis values indicate rater angle measure was greater than electrogoniometer criterion reference angle measure. deg, degrees.

Intrarater Reliability

The $ICC_{(3,1)}$ values calculated using rater measures from both testing sessions ranged from 0.994 to 0.998 (Table 2). Additionally, intrarater reliability SEM values ranged from 0.43 to 0.84 degrees (Table 2). Based on the largest SEM (0.84), 95% of errors upon retesting are expected to be at most 1.64 degrees.

Table 2. Intrarater Reliability

	Rater 1	**Rater 2**	**Rater 3**	**Rater 4**	**Rater 5**	**Rater 6**
$ICC_{(3,1)}$ (95%CI)	0.997 (0.993-0.999)	0.994 (0.982-0.998)	0.996 (0.989-0.999)	0.998 (0.995-0.999)	0.997 (0.992-0.999)	0.994 (0.983-0.998)
SEM (°)	0.53	0.84	0.66	0.44	0.57	0.83
CI, confidence interval; SEM, standard error of measurement						

Interrater Reliability

The $ICC_{(2,1)}$ and SEM for interrater reliability were 0.996 and 0.67 degrees, respectively. Based on these results, 95% of measurement errors between raters are expected to be 1.32 (1.96 x SEM) degrees or less.

DISCUSSION

The current study examined criterion validity and reliability of knee FPPA during a single-limb stepdown task measured using the Technique application and found it to be reliable, valid, and therefore suitable for clinical use. This was determined through criterion validity and inter- and intrarater reliability which was established for knee FPPA measures made using the Technique application on an iPad, stylus, and with skin markers. Regarding validity, an important consideration is whether the Technique application, as utilized in the current study, is sufficiently close to our criterion validity of an electrogoniometer to detect clinically relevant differences. For example, it has been reported that PFP subjects demonstrate 2.7 degrees (41%) greater coronal plane knee motion during a stepdown.[11] This is further supported by Willson et al. who showed a 4.1 degree greater knee FPPA in individuals with PFP compared to asymptomatic subjects.[10] The MAD between rater measures obtained using Technique and the electrogoniometer measures (for all raters) was less than one degree. Additionally, based on the largest LA observed, 95% of the differences between rater measures of the knee FPPA and the true knee FPPA would be 2.04 degrees or less. Based on these findings it appears the Technique application is useful for detecting clinically relevant knee FPPA differences.

Concerning reliability, the findings from the current study are not consistent with some previously reported results. For example, in two prior studies, poor to moderate interrater reliability ICC values were reported when using 2D video and an application to quantify the FPPA at the knee (ICC=0.51)[15]and hip (ICC=0.31).[16] In contrast, the interrater reliability value observed in the current study was substantially greater (ICC=0.996). Additionally, the SEM value pertaining to interrater reliability in the present study (0.67 degrees) was approximately one-tenth of the SEM reported by King and Belyea (7.2 degrees).[15] These differences are likely attributable to the additional features used in the current study to improve precision such as skin markers and a stylus for aligning the angle measurement tool.

Additionally, the tablet used in this study had a 32.8 cm screen compared to the 24.6 cm and 11.9 cm screens used by King and Belyea[15] and Neal et al.[16], respectively. The bigger screen may have further enhanced precision in aligning the angle measurement tool in the application.

This study has several limitations. First, measurement errors caused by non-planar movement were not accounted for. Additionally, skin marker placement reliability was not examined in the current study. It is unknown if or to what extent marker placement error at subsequent testing sessions wiykd influence the usefulness of Technique. Finally, the skin markers used in this study were placed on soft tissue to represent underlying bones. Measurement error during single-limb stepdown caused by skin movement over underlying bone was not accounted for and could have introduced error.

Future studies should examine the effects of marker placement issues and movement artefact on knee FPPA measurement error when using Technique. Additionally, it would be worthwhile to explore whether knee FPPA measures made using Technique predict PFP development.

CONCLUSION

These findings indicate knee FPPA estimates obtained using Technique on an iPad, with stylus and skin markers have excellent reliability (intrarater 1.64 degrees, interrater 1.32 degrees) and satisfactory agreement with criterion references (2.04 degrees or less). Furthermore, Technique is suitable for broad-based clinical implementation and is of value to clinicians for objectively documenting movement dysfunction and treatment effectiveness.

ACKNOWLEDGEMENTS

This study was funded by an Instructional Innovation and Research Development Mini Grant from Angelo State University.

REFERENCES

1. Smith BE, Selfe J, Thacker D, et al. Incidence and prevalence of patellofemoral pain: A systematic review and meta-analysis. Screen HR, ed. *PLOS ONE*. 2018;13(1):e0190892. doi:10.1371/journal.pone.0190892

2. Huberti HH, Hayes WC. Patellofemoral contact pressures. the influence of q-angle on tendofemoral contact. *J Bone Jt Surg*. 1984;66(5):715-724.

3. Powers CM, Witvrouw E, Davis IS, Crossley KM. Evidence-based framework for a pathomechanical model of patellofemoral pain: 2017 patellofemoral pain consensus statement from the 4th International Patellofemoral Pain Research Retreat, Manchester, UK: part 3. *Br J Sports Med*. 2017;51(24):1713-1723. doi:10.1136/bjsports-2017-098717

4. Nakagawa TH, Moriya ETU, Maciel CD, Serrao FV. Frontal plane biomechanics in males and females with and without patellofemoral pain. *Med Sci Sports Exerc*. 2012;44(9):1747-1755. doi:10.1249/MSS.0b013e318256903a

5. Herrington L. Knee valgus angle during single leg squat and landing in patellofemoral pain patients and controls. *The Knee*. Published online November 27, 2013. doi:10.1016/j.knee.2013.11.011

6. Holden S., Boreham C., Doherty C., Delahunt E. Two-dimensional knee valgus displacement as a predictor of patellofemoral pain in adolescent females. *Scand J Med Sci Sports*. 2015;27(2):188-194. doi:10.1111/sms.12633

7. Almeida GPL, Carvalho e Silva AP de MC, França FJR, Magalhães MO, Burke TN, Marques AP. Does anterior knee pain severity and function relate to the frontal plane projection angle and trunk and hip strength in women with patellofemoral pain? *J Bodyw Mov Ther*. 2015;19(3):558-564. doi:10.1016/j.jbmt.2015.01.004

8. Almeida GPL, Silva AP de MCC e, França FJR, Magalhães MO, Burke TN, Marques AP. Relationship between frontal plane projection angle of the knee and hip and trunk strength in women with and without patellofemoral pain. *J Back Musculoskelet Rehabil*. 2016;29(2):259-266. doi:10.3233/BMR-150622

9. Emamvirdi M, Letafatkar A, Khaleghi Tazji M. The effect of valgus control instruction exercises on pain, strength, and functionality in active females with patellofemoral pain syndrome. *Sports Health*. 2019;11(3):223-237. doi:10.1177/1941738119837622

10. Willson JD, Davis IS. Utility of the frontal plane projection angle in females with patellofemoral pain. *J Orthop Sports Phys Ther*. 2008;38(10):606-615. doi:10.2519/jospt.2008.2706

11. Richards J, Selfe J, Miscp SK, Callaghan M, Atkins L. Are there differences in knee stability between patients with patellofemoral pain and healthy subjects during a slow step descent task? 2018;(3):4.

12. Lopes Ferreira C, Barton G, Delgado Borges L, dos Anjos Rabelo ND, Politti F, Garcia Lucareli PR. Step down tests are the tasks that most differentiate the kinematics of women with patellofemoral

pain compared to asymptomatic controls. *Gait Posture*. 2019;72:129-134. doi:10.1016/j.gaitpost.2019.05.023

13. Whittle's Gait Analysis - 5th Edition. Accessed April 1, 2021. https://www.elsevier.com/books/whittles-gait-analysis/levine/978-0-7020-4265-2

14. Atkins LT, James CR, Sizer PS, Jonely H, Brismée JM. Reliability and concurrent criterion validity of a novel technique for analyzing hip kinematics during running. *Physiother Theory Pract*. 2014;30(3):210-217. doi:10.3109/09593985.2013.830349

15. King DL, Belyea BC. Reliability of using a handheld tablet and application to measure lower-extremity alignment angles. *J Sport Rehabil*. 2015;24(4). doi:10.1123/jsr.2014-0195

16. Neal BS, Lack SD, Barton CJ, Birn-Jeffery A, Miller S, Morrissey D. Is markerless, smart phone recorded two-dimensional video a clinically useful measure of relevant lower limb kinematics in runners with patellofemoral pain? A validity and reliability study. *Phys Ther Sport*. 2020;43:36-42. doi:10.1016/j.ptsp.2020.02.004

17. Payton CJ, Bartlett RM, British Association of Sport and Exercise Sciences, eds. *Biomechanical Evaluation of Movement in Sport and Exercise: The British Association of Sport and Exercise Sciences Guidelines*. 1. publ. Routledge; 2008.

18. Indramohan VP, Valsan G, Rowe PJ. Development and validation of a user-friendly data logger (SUDALS) for use with flexible electrogoniometers to measure joint movement in clinical trials. *J Med Eng Technol*. 2009;33(8):650-655. doi:10.3109/03091900903159403

19. Martin Bland J, Altman DouglasG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;327(8476):307-310. doi:10.1016/S0140-6736(86)90837-8

20. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3 edition. Prentice Hall; 2009.