Journal of Sport Rehabilitation, 2026, 35, 271-291
https://doi.org/10.1123/jsr.2025-0162

Human Kinetics
SYSTEMATIC REVIEW

Quantitative Measurement of Rotational Knee Stability: A Systematic Review of Instrument Reliability and Validity

Frédérique A.C. Oosterbaan,[1] Mariska G.H. Wesseling,[1,2] Tom M. Piscaer,[2] Tjep Hoedemakers,[2] Jaap Harlaar,[1,2] and Erin M. Macri[2]

[1]BioMechanical Engineering, Delft University of Technology, Delft, The Netherlands; [2]Department of Orthopaedics and Sports Medicine, Erasmus MC University Medical Centre Rotterdam, Rotterdam, The Netherlands

Context: Objectively evaluating knee stability in multiple planes in individuals with anterior cruciate ligament injury may provide more comprehensive information than evaluating subjectively or in only a single plane. This could support both research and clinical decision making. However, for the clinical value of such an instrument to be evaluated, reliability and validity of the instrument must first be established. Despite multiple available instruments that measure rotational knee stability, it is not clear which of these instruments has adequate reliability and validity. ***Objective:*** We performed a systematic review to identify instruments for measuring rotational knee stability and to synthesize the available literature in which validity and reliability were evaluated. ***Evidence Acquisition:*** We searched 4 databases for publications reporting reliability or validity of an instrument designed to assess rotational knee stability. A narrative synthesis was used to present the results. ***Evidence Synthesis:*** We identified 42 studies evaluating 25 different instruments designed to measure movement while applying a standardized torque or while a tester performed a manual test (eg, pivot shift). There was high heterogeneity in parameters reported and criterion methods used. Intrarater and interrater reliability intraclass correlation coefficients were consistently adequate (>.75) except for when lower torques (ie, 6 N·m or less) were applied or acceleration or jerk was measured instead of laxity. Four out of 19 (21.1%) studies evaluating validity reported very good correlations ($r > .8$) with a criterion measure. ***Conclusions:*** We found no high-quality evidence that provided sufficient evidence of both reliability and validity in any device. To evaluate the clinical benefit of objectively evaluating stability in multiple planes, further work is needed to develop, refine, and evaluate this class of devices.

Keywords: knee laxity, measurement properties, anterior cruciate ligament, medical devices

Key Points

- Evaluating knee stability in both anteroposterior and rotational directions could yield a more comprehensive assessment of knee stability than anteroposterior alone.
- Methods matter: Reliability of devices measuring rotational knee stability is generally adequate, but lower torques and measures of acceleration or jerk may be less reliable.
- No devices that measure rotational knee stability have been sufficiently evaluated for both reliability and validity.

Context

The anterior cruciate ligament (ACL) is the most frequently injured knee structure, with approximately 80,000 noncontact injuries in the United States each year.[1] ACL tears result in decreased joint stability through increased joint laxity, that is, increased anterior translation and internal rotation of the tibia relative to the femur in response to imposed passive joint loads.[2]

Quantifying knee stability after injury is necessary to inform clinical decision making. Knee stability is generally assessed by performing clinical tests, such as the pivot-shift test,[3] or with instruments that apply standardized forces to move the tibia relative to the femur, such as the KT-1000 (MedMetric).[4,5] Clinical tests depend on the performance and interpretation of the examiner, making them more subjective than instrumented tests.[6] The validity and reliability of instruments have been widely evaluated to assess stability in the anteroposterior direction.[6,7] However, injury mechanisms involve excessive rotational load, and knee rotation is a major component of knee stability. Therefore, further evaluating stability in a rotational direction would yield a more comprehensive assessment of knee stability. This could enhance clinical decision making, for example, to determine whether it is indicated to refer to a specialist or for additional imaging, or to assist with surgical planning. Unfortunately, reliability and validity have not been fully established for the rotational direction.[8–10] Therefore, it remains unknown which device is best for objectively evaluating rotational stability and, consequently, to what extent such a device would improve clinical care.

We, therefore, performed a systematic review to synthesize the available literature in which validity and reliability were evaluated in instruments that objectively measure rotational knee stability.

Wesseling https://orcid.org/0000-0002-3347-2130
Piscaer https://orcid.org/0000-0002-9448-9948
Hoedemakers https://orcid.org/0009-0005-5590-6517
Harlaar https://orcid.org/0000-0003-2889-271X
Macri (e.macri@erasmusmc.nl) is corresponding author, https://orcid.org/0000-0003-2798-6052

Methods

Design

We prospectively registered our protocol with PROSPERO (#CRD42022308643) in March 2022. We adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to report this systematic review.[11]

Evidence Acquisition

Search Strategy

We searched 4 digital databases: PubMed, Embase, CINAHL, and SPORTDiscus. Keywords and medical subject headings (MeSH terms) were used to identify eligible studies (see Appendix A: Search terms). We performed the initial search in January 2022 and updated it in July 2024. No filters or limits were used.

Eligibility Criteria

We included studies in which reliability or validity was evaluated in instruments that assessed passive rotational stability of the tibiofemoral joint in adolescents or adults with ACL deficiency, ACL reconstruction or repair, or intact ligaments. We included studies in which concomitant injuries (eg, meniscus tears) were present along with the ACL injury. There was no restriction on the criterion method used. The instruments were either designed to objectively measure movement (ie, internal/external rotation or related measurements, like linear acceleration during rotation) while an examiner performed manual testing (eg, pivot-shift test), or they were designed to apply a standardized torque and measure the resultant movement. We included any study design in which validity or reliability could be assessed, such as observational study designs or randomized controlled trials. We included studies published in English, Dutch, or French.

We excluded studies that evaluated patellofemoral join stability or that evaluated an instrument's validity or reliability in an anteroposterior direction only, that is, without assessing rotational stability. We also excluded studies using invasive instrumentation or instruments evaluated during surgery, though preoperative evaluations under anesthesia were included.

Selection Process

We imported all records into EndNote 21 to remove duplicates.[12,13] The remaining records were then uploaded to an online screening platform, Rayyan (https://www.rayyan.ai/) during the first search and Covidence (https://www.covidence.org/) during the updated search. Two independent reviewers performed the title and abstract screening and subsequent full-text screening. Disagreements were discussed with a third, senior reviewer. Finally, we screened the reference lists of all included studies, alongside those of relevant (systematic) review papers, for additional possible records.

Data Extraction

Data collection was performed independently by 2 reviewers. Any uncertainties were discussed between the reviewers and senior coauthors as needed. Data extracted included the publication (first author, journal, and year of publication), the instrument used, participant demographics (sex, age, and body mass index), study design, and measurement properties (reliability and validity). Measurement properties were only extracted for internal/external rotation or related measurements (eg, linear acceleration during rotation). We defined measurement properties (ie, reliability and validity) according to the COSMIN taxonomy.[14]

For reliability, we extracted intrarater and interrater reliability intraclass correlation coefficients (ICCs) along with the standard error of measurement (SEM) when available. Otherwise, we extracted the reliability measures reported in the study (eg, coefficient of variation). If reliability was evaluated multiple times under similar conditions (eg, different knee-flexion angles), we only extracted the most conservative value. ICCs above .75 were considered *adequate*.[15]

For validation, we extracted results for both continuous and dichotomous outcomes. For continuous outcomes, we extracted results for concurrent validation and construct validation (ie, known groups, convergent, and divergent validity) and reported Pearson correlation coefficients (r) or the most closely related statistic (eg, Spearman correlations). Correlations above .8 were considered *very good*, .6 to .8 were *moderately strong*, .3 to .6 were *fair*, and below .3 were *poor*.[16] For dichotomous outcomes, we extracted sensitivity, specificity and the criterion method used for comparison (ie, criterion defining presence of ACL injury). In addition, the agreement, area under the receiver operating curve (AUC), and cutoff values were extracted when reported. AUCs above 0.7 were considered *acceptable*, above 0.8 *excellent*, and above 0.9 *outstanding*.[17]

Risk of Bias Assessment

To assess the risk of bias, we used the COSMIN criteria[18] for reliability (box 6) and validity (box 8) with continuous outcomes and QUADAS 2[19] for validity with dichotomous outcomes. Box 6 of the COSMIN checklist contains 3 domains: design requirements, statistical methods, and other.[18] Box 8 of the checklist contains 2 domains: statistical methods and other. Each domain contains one or more questions, with response options of *very good, adequate, doubtful,* or *inadequate*. The overall risk of bias for a study corresponds to the item with the lowest score. A *very good* score is rated as low risk of bias; *adequate* scores have relatively low risk of bias; *doubtful* scores have relatively high risk of bias; and *inadequate* scores are rated as high risk of bias. QUADAS 2 contains 4 domains: patient selection, reference standard, index test, and flow and timing.[19] Each domain contains one or more questions, and response options are *yes, no,* or *unclear*. Like COSMIN, QUADAS 2 yields overall scores of *low, high*, or *unclear* risk of bias based on the lowest score. Each study was assessed independently by 2 reviewers, and uncertainties were discussed among the reviewers and senior coauthors.

Protocol Deviations

Initially, our main aim was to evaluate the instrumented measurement of rotational knee stability. However, as we expected to find few papers on this subject, we utilized a broad search strategy to include papers about instruments measuring stability in all directions and in both tibiofemoral and patellofemoral joints. Given the large number of available papers identified, we were able to narrow the scope of our review to focus on tibiofemoral rotational stability in individuals with or without a history of ACL tear. In the protocol, we also stated we would include meta-analyses and (systematic) reviews. However, the identified reviews focused on different stability directions or conditions (eg, intraoperative); thus, we only included original studies of rotational stability. We did, however, review reference lists of

identified reviews in search of records that may have been missed with our search strategy.

Synthesis of Results

Due to study heterogeneity among the included studies (eg, differences in methods and statistics used to measure validity and reliability), meta-analysis was not possible. Therefore, we performed a narrative synthesis.

Evidence Synthesis

Study Selection

A total of 4395 records were retrieved, from which 1037 duplicates were removed. Based on the title and abstract screening, 3108 records were excluded. Following full-text screening and review of reference lists, we ultimately included 42 papers (see Figure 1, Table 1).

Study Characteristics

Reliability was evaluated in 30 (71.4%) studies, and validity (including diagnostic accuracy) was evaluated in 23 (54.8%) studies. Sample sizes ranged from 1 to 295 participants but were fewer than 50 in 30 (71.4%) studies. Most studies included both men and women, but 6 (14.3%) included only men, and 5 (11.9%) did not fully report sex. Mean ages, when reported, ranged from 21 to 36 years old. Mean body mass index ranged from 22.0 to 25.9 kg/m^2 but was only fully reported in 10 studies (23.8%). Seventeen (40.5%) studies evaluated intact knees only, 10 (23.8%) evaluated ACL-deficient knees only, 2 (4.8%) studies evaluated ACLR knees only, and the remainder (31.0%) evaluated multiple groups (7 ACL deficient and intact knees, 3 ACL-deficient and ACLR knees, 2 ACLR and intact knees, 1 all 3). Nine (21.4%) studies evaluated patients under anesthesia to reduce voluntary muscle resistance to testing.

Instruments

We identified 25 different instrumented setups (see Table 2). Six (14.3%) studies evaluated stability using electromagnetic sensors during a pivot-shift test. Eight (19.1%) studies used magnetic resonance imaging (MRI) with various instruments or manual tests used in the MRI. Sixteen devices were evaluated in a single study only.

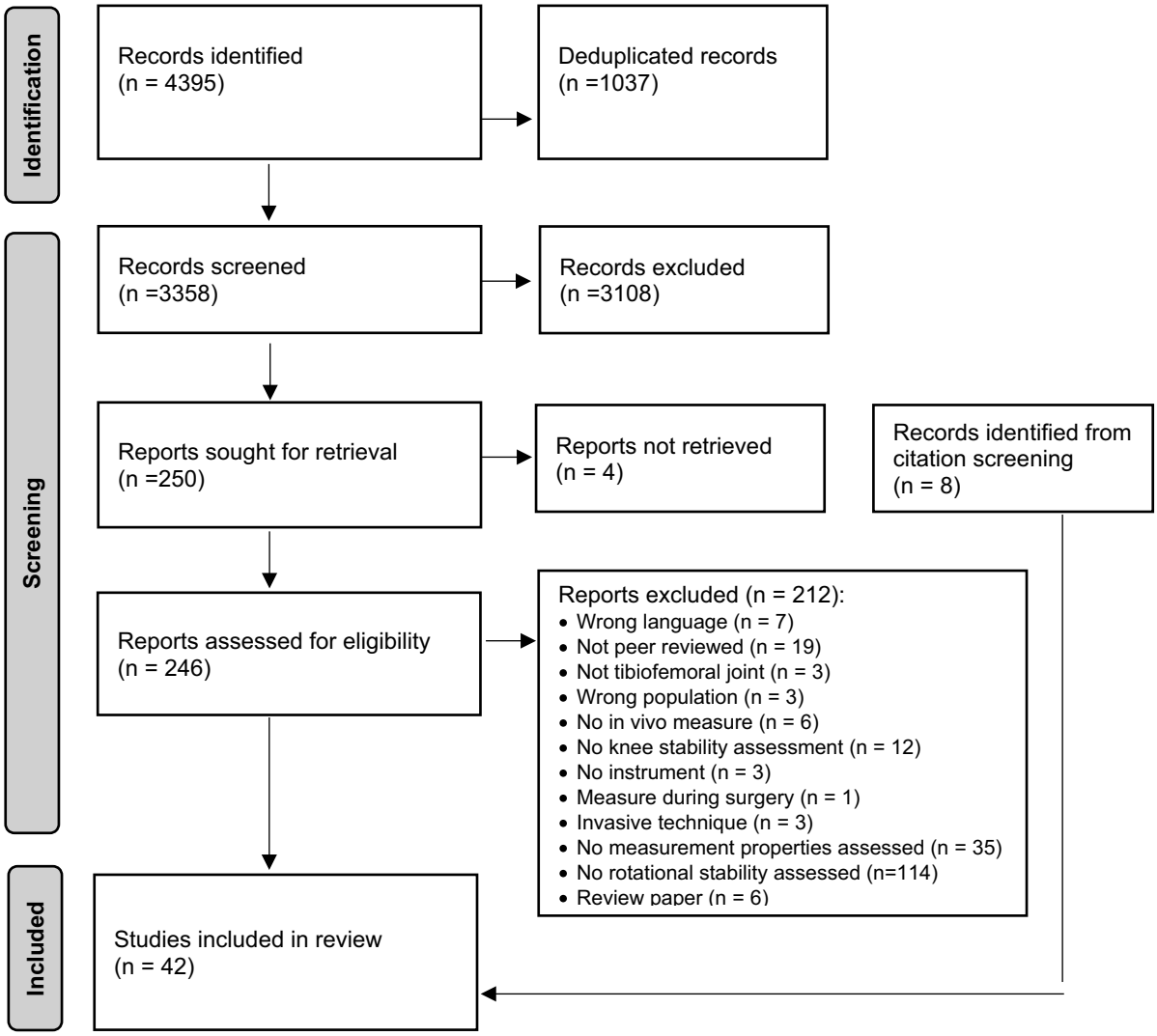

Figure 1 — PRISMA flowchart.

Table 1 Characteristics of the Included Studies

Study	Instrument used	Sample size (n)	Sex (n)	Age (mean [SD], y)	BMI (mean [SD], m/kg^2)	Knee condition
Okazaki et al[20]	Horizontally open-bore MRI during Slocum test	50	G 1: 10m/8f G 2: 9m/9f G 3: 8m/6f	G 1: 29.3 (7.7) G 2: 29.5 (7.8) G 3: 26.3 (6.8)	NR	G 1: unilateral ACLD with 4 failed reconstructions G 2: intact G 3: unilateral ACLD
Tashiro et al[50]	Horizontally open-bore MRI during Slocum test	41	G 1: 9m/12f G 2: 9m/11f	G 1: 24 (r: 15–47) G 2: 25 (r: 15–41)	NR	G 1: unilateral ACLR G 2: unilateral ACLD
Okazaki et al[51]	Horizontally open-bore MRI during Slocum test	92	G 1: 17m/19f G 2: 9m G 3: 22f G 4: 8m/17f	G 1: 27.2 G 2: 28.4 G 3: 26.1 G 4: 27.0	NR	G 1: chronic ACLD G 2–4: ACLR
Tardy et al[21]	Horizontally open-bore MRI with KneeM device	10	7m/3f	32.3 (9.4)	NR	Unilateral ACLD without meniscal pathology
Espregueira-Mendes et al[52]	MRI with PKTD	61	61m	29 (9.3)	24.9 (2.6)	Unilateral ACLD, 30% also had a meniscus injury. In 16 participants, the intact contralateral knee was also tested.
Hemmerich et al[22]	Horizontally open-bore MRI with torsional loading applied manually via pulling a cord attached to an MRI-mounted boot device	6	5m/1f	29.3 (3.6)	Height: 178.0 (8.6) cm Mass: 72.0 (13.0) kg	Intact
Carpenter et al[23]	MRI with torsional and axial loading applied with weights in a pulley system integrated in a boot rig	14 (24 knees)	3m/11f	32 (10)	NR	Intact
Haughom et al[24]	MRI with torsional and axial loading applied with weights in a pulley system integrated in a boot rig	26	G 1: 8m/8f G 2: 5m/5f	G 1, m: 26.5 (6.7) G 1, f: 27.0 (3.9) G 2, m: 33.8 (10.4) G 2, f: 36.3 (10.7)	G 1, m: 24.0 (1.9) G 1, f: 22.0 (1.3) G 2, m: 24.1 (1.8) G 2, f: 22.0 (1.7)	G 1: intact G 2: ACLD
Hoshino et al[25]	Electromagnetic measurement during pivot shift	30	14m/16f	21.2 (r: 14–39)	NR	Unilateral ACLR Evaluated under anesthesia
Kubo et al[26]	Electromagnetic measurement during pivot shift	G 1: 5 G 2: 25	G 1: NR G2: 11m/14f	G 1: NR G 2: 26.3 (r: 15–34)	NR	Unilateral isolated ACLD G 1 was evaluated under anesthesia.
Labbe et al[46]	Electromagnetic measurement during pivot shift	65	G 1: 9m/3f G 2: 32m/26f	G 1: 32.9 G 2: 26.8	NR	G 1: 4 intact and 8 ACLD G 2: 8 intact and 45 ACLD
Kuroda et al[27]	Electromagnetic measurement during pivot shift	1	m	33	NR	Unilateral ACLD Evaluated under anesthesia
Araki et al[60]	Electromagnetic measurement during pivot shift	G 1: 20 G 2: 20	G 1: 10m/10f G 2: 10m/10f	G 1: 27.6 (10.5) G 2: 24.7 (10.1)	NR	G 1: partial ACLD with 2 medial and 3 lateral meniscal tears G 2: complete ACLD with 8 medial and 5 lateral meniscal tears Both groups were evaluated under anesthesia.
Labbe et al[45]	Support vector machine algorithm based on electromagnetic measurement during pivot shift	56	G 1: 9m/3f G 2: 26m/18f	G 1: 32.9 (9.2) G 2: 26.6 (11.5)	NR	G 1: 4 intact, 4 ACLD G 2: 4 intact, 40 ACLD

(continued)

Table 1 *(continued)*

Study	Instrument used	Sample size (n)	Sex (n)	Age (mean [SD], y)	BMI (mean [SD], m/kg^2)	Knee condition
Shultz et al[28]	Electromagnetic measurement with VKLD	20	10m/10f	m: 27.3 (3.4) f: 22.9 (1.5)	m: 25.9 f: 23.1	Intact
Branch et al[47]	Electromagnetic measurement with RKT device	93	G 1: 10m/4f G 2: 51m/28f	G 1: 36.0 (11.6) G 2: 34.2 (11.4)	G 1: 23.5 (2.3) G 2: 24.4 (3.1)	G 1: intact G 2: ACLR
Beckley et al[29]	Electromagnetic measurement with RKT device	25	NR	r: 20–48	NR	Intact
Moewis et al[53]	Optical motion capture system with Knee Joint Laxity Testing Device	4	3m/1f	34 (15)	24 (3)	Unilateral ACLD
Tsai et al[30]	Magnetic motion capture system with Knee Laxity Measurement Device	11	11m	30.3 (r: 27–35)	25.6	Intact
Mayr et al[54]	Polymer-based sensor elastic–capacitive measurement system, torque applied manually through a boot	10	5m/5f	28 (6)	NR	Intact
Bellitti et al[55]	Smart Brace Design using IMUs during pivot shift	4	NR	NR	NR	Suspected unilateral ACLD, both knees tested for all patients
Lopomo et al[31]	KiRA during pivot shift	51	40m/11f	30.9 (r: 16–63)	NR	Unilateral ACLD, both knees tested Evaluated under anesthesia
Berruto et al[32]	KiRA during pivot shift	100	65 m/35f	29 (9)	NR	Unilateral ACLD, 46 with additional meniscal tears
Nakamura et al[59]	KiRA during pivot shift	29	17m/12f	24 (range 14–46)	NR	Unilateral primary double-bundle ACLR with an autologous semitendinosus tendon. Seven patients had both medial and lateral meniscus injuries that were repaired; 5 had a medial meniscus injury, 4 were repaired, and 1 was partially removed. Seven patients had a lateral meniscus injury that was repaired. Both knees were tested with patient awake and under anesthesia.
Helfer et al[56]	KiRA during pivot shift	295	233m/62f	24.8 (4.5)	23.8 (2.6)	Unilateral ACLD Evaluated under anesthesia
Napier et al[48]	KiRA during pivot shift	50	26m/24f	28.6	NR	Unilateral ACLD Evaluated under anesthesia
Vaidya et al[33]	Smartphone accelerometer during pivot shift	17	14m/3f	33 (12)	NR	Unilateral ACLD with 2 lateral meniscal injuries, 5 medial meniscal injuries, 2 partial medial collateral ligament injuries, and 1 previously reconstructed ACL. Evaluated under anesthesia
Krause et al[44]	Handheld inclinometer during dial test	24	11m/13f	22.5 (2.8)	25.7	Intact
Alam et al[34]	RMD with inclinometer during the dial test	46	25m/21f	27 (r: 21–34)	No obese participants	Intact
McQuade et al[35]	Genucom knee analysis system	5	3m/2f	r: 20–25	NR	Intact
Wroble et al[49]	Genucom knee analysis system	14	G 1: 5m/5f G 2: 1f G 3: NR	G 1: 23 G 2: 24 G 3: NR	NR	G 1: intact G 2: intact G 3: unilateral chronic ACLD

(continued)

Table 1 *(continued)*

Study	Instrument used	Sample size (n)	Sex (n)	Age (mean [SD], y)	BMI (mean [SD], m/kg^2)	Knee condition
Mills and Hull[36]	Rotatometer	3	m	29.7 (8.0)	Height: 1.77 (0.04) m Weight: 73.3 (5.68) kg	Intact
Chung et al[37]	Rotatometer	G 1: 53 G 2: 41	G 1: 33m/20f G 2: 21m/20f	G 1: 31 (9) G 2: 24 (3)	G 1: 24 (2) G 2: 24 (3)	Intact
Lorbach et al[38]	Rotameter	30	15m/15f	24.1	Height: 1.75 (0.10) m Weight: 69 (13) kg	Intact
Mouton et al[61]	Rotameter	128	89m/39f	m: 28 (9) f: 27 (11)	Height, m: 1.79 (0.07) m Weight, m: 80 (12) kg Height, f: 1.68 (0.07) m Weight, f: 67 (10) kg	Intact
Neumann et al[39]	Rotameter, updated prototype	1	NR	NR	NR	Intact
Lorenz et al[40]	KUKA robot	40	20m/20f	29 (13)	NR	Intact
Almquist et al[57]	Rottometer	5	m	NR	NR	Previous ACLR
Almquist et al[41]	Rottometer	20	G 1: 4m/6f G 2: 4m/6f	G 1: r: 28–69 G 2: r: 24–46	NR	Intact
Mayr et al[58]	Laxitester	G 1: 24 G 2: 23	G 1: 14m/10f G 2: 13m/10f	G 1: 31.3 (9.1) G 2: 31.2 (6.7)	18–30	G 1 with primary unilateral ACLD ± concomitant meniscus injuries G 2 ACL rerupture ± concomitant meniscus injuries
Nascimento et al[42]	DYNEELAX	48	34m/14f	35.8 (10.5)	25.4 (4.1 9)	82 intact, 11 ACLR, and 3 ACLD. Both knees tested for all patients, one was only tested for anterior translation.
Kang et al[43]	Computational modeling derived from the Telos stress test	1	m	35	NR	Intact

Abbreviations: ACL, anterior cruciate ligament; ACLD, anterior cruciate ligament deficient; ACLR, anterior cruciate ligament reconstructed; BMI, body mass index; f, female; G, group; IMUs, inertial measurement units; KiRA, Kinematic Rapid Assessment; MRI, magnetic resonance imaging; m, male; NR, not reported; PKTD, Porto knee testing device; r, range; RKT, robotic knee testing; RMD, rotational measurement device; VKLD, Vermont knee laxity device.

Table 2 Instruments Used to Evaluate Rotational Knee Stability

Study	Force/torque application or movement	Measurement technique
Okazaki et al[20] Okazaki et al[51] Tashiro et al[50]	Manual Slocum test	Open-bore MRI images
Tardy et al[21]	KneeM device; applies an anterior draw and internal rotation to the lower-limb with the aid of a dynamometer	Open-bore MRI images
Espregueira-Mendes et al[52]	PKTD; mechanically applies either an internal or external rotation at the foot	MRI images
Hemmerich et al[22]	Device to mechanically apply torsional loading at the foot by pulling on a cord attached to the device	Open-bore MRI images
Carpenter et al[23] Haughom et al[24]	Device to mechanically apply torsional and axial loading at the foot	MRI images
Araki et al[60] Hoshino et al[25] Kubo et al[26] Kuroda et al[27] Labbe et al[46]	Manual pivot-shift test	Electromagnetic measurements
Labbe et al[45]	Manual pivot-shift test	Support vector machine algorithm with input from an electromagnetic measurement
Shultz et al[28]	VKLD; counteracts gravity with weights and, thereby, the shear load in the tibiofemoral joint. Rotations are then applied using a T-handle and force transducer at the foot	Electromagnetic measurements
Beckley et al[29] Branch et al[47]	RKT device; mechanically applies rotations at the foot	Electromagnetic measurements
Moewis et al[53]	Knee joint laxity testing device; applies axial rotations with the aid of a force transducer at the foot	Optical motion capture system
Tsai et al[30]	Knee laxity measuring device; applies rotations at the foot through a handlebar with a sensor	Magnetic motion tracking
Mayr et al[54]	Torque manually applied at the foot through a boot	Custom-made polymer-based sensor elastic–capacitive measurement system
Bellitti et al[55]	Manual pivot-shift test	Smart brace design consisting of 3 stretch sensors and 2 IMUs
Berruto et al[32] Nakamura et al[59] Helfer et al[56] Lopomo et al[31] Napier et al[48]	Manual pivot shift	KiRA: measures the linear accelerations during the pivot-shift phenomenon
Vaidya et al[33]	Manual pivot shift	Smartphone accelerometer: measures the peak linear accelerations during the pivot shift phenomenon
Krause et al[44]	Manual dial test	Handheld inclinometer
Alam et al[34]	Boot where torques can be applied to the foot using a torque wrench	RMD; uses inclinometers at the femur and tibia
McQuade et al[35] Wroble et al[49]	Manual stability tests	Genucom: electrogoniometers
Chung et al[37] Mills and Hull[36]	Rotameter: torque manually or mechanically applied to the foot	Rotameter: electrogoniometers
Lorbach et al[38] Mouton et al[61] Neumann et al[39]	Rotameter: torques are applied at the foot through a handlebar and electronic torque key	Rotameter: electrogoniometers
Lorenz et al[40]	KUKA robot: force or torque is applied at the shank	KUKA robot: kinematics of the robot during testing
Almquist et al[57] Almquist et al[41]	Rottometer: torque is applied using an adjustable spanner	Rottometer: measuring stick at the foot
Mayr et al[58]	Laxitester: torque mechanically applied to the foot plate	Laxitester: measures at the foot
Nascimento et al[42]	DYNEELAX: mechanically apply torque at the foot	DYNEELAX; displacement transducer
Kang et al[43]	Stress radiography using the Telos device to apply an anterior or posterior loading	Computational modeling

Abbreviations: IMU, inertial measurement unit; KiRA, Kinematic Rapid Assessment; MRI, magnetic resonance imaging; PKTD, Porto knee testing device; RKT, robotic knee testing; RMD, rotational measurement device; VKLD, Vermont knee laxity device.

Table 3 Reliability

Study	Instrument used	Intrarater/test–retest	Interrater
Okazaki et al[20]	Horizontally open-bore MRI during Slocum test	ICC: .96–.98 ($P < .01$) Difference between tests: 1.0 (SD 0.7) mm	Full procedure ICC: .91 ($P = .02$) Difference between tests: 1.1 (SD 0.7) mm Image measure only ICC: .93 ($P < .01$) Difference between tests: 1.8 (0.3) mm
Tardy et al[21]	Horizontally open-bore MRI with KneeM device	ICC: - Intact: ≥.85 - ACLD: ≥.93	ICC: - Intact: ≥.82 - ACLD: ≥.92
Hemmerich et al[22]	Horizontally open-bore MRI with torsional loading through pulling a cord	ICC, torque normalized to body mass: - Extended: .99 - Flexed: .93	
Carpenter et al[23]	MRI with torsional and axial loading applied with weights in a pulley system	RMS average of test–retest precision errors, IR at 2.7 N·m torque: 1.6°	
Haughom et al[24]	MRI with torsional and axial loading applied with weights in a pulley system	ICC(2,1)/SEM, at 3.35 N·m torque: .91/1.1°	
Hoshino et al[25]	Electromagnetic measurement during PS	Correlation coefficient matrix: - Coupled anterior tibial translation: .98–.99 ($P < .05$, SD: 1.1 [0.6] mm) - Acceleration of posterior translation: .95–.97 ($P < .05$, SD: –211 [176] $\frac{mm}{s^2}$)	
Kubo et al[26]	Electromagnetic measurement during PS	CoV (mean of 5 trials): - Lateral translation: 30% - Posterior translation: 23%	
Labbe et al[46]	Electromagnetic measurement during PS		ICC for maximum linear acceleration, before/after standardization: .41/.52
Kuroda et al[27]	Electromagnetic measurement during PS	ICC (3 trials): - Tibial IR: .98 - Acceleration: .87	
Labbe et al[45]	Support vector machine algorithm based on electromagnetic measurement during PS		$K = 0.81$
Shultz et al[28]	Electromagnetic measurement with VKLD	ICC(2,*k*)/SEM, at 0–5 N·m torque: Unloaded: - IR: .89/1.56° - ER: .86/2.99° - Total IR + ER: .89/2.80° Compressive load of 40% body weight, at 0–5 N·m torque: - IR: .20/1.70° - ER: .85/1.45° - Total IR + ER: .70/2.64°	
Branch et al[47]	Electromagnetic measurement with RKT device		ICC, at 5.65 N·m torque: - Total IR/ER: .97 - Compliance[a]: .97
Beckley et al[29]	Electromagnetic measurement withRKT device	ICC(2,1)/SEM, at 6 N·m torque: ≥.83/0.14°–0.18°	
Tsai et al[30]	Magnetic motion capture system with knee laxity measurement device	Intrarater ICC(2,1)/SEM, at 6 N·m torque (95% CI): - 30° and 90° flexion: >.95/<1° (<2°) Test–retest ICC(2,5)/SEM, at 2, 4 and 6 N·m torque (95% CI): - 90° flexion: .77/1.9° (3.8°)	ICC(2,5)/SEM, at 6 N·m torque (95% CI): - 30° flexion: .81/2.6° (5.1°) - 90° flexion: .88/1.6° (3.2°)
Lopomo et al[31]	KiRA during PS	ICC (95% CI): - ACLD, minimal acceleration: .75 (.58 to .86) - ACLD, maximal acceleration: .93 (.85 to .96) - Intact, minimal acceleration: .69 (.42 to .84) - Intact, maximal acceleration: .76 (.55 to .88)	
Berruto et al[32]	KiRA during PS	Correlation coefficient: .7–.9 ($P < .05$) There is a learning curve required to equalize data	
Napier et al[48]	KiRA during PS		ICC(3,1): NR Spearman correlation between 3 surgeons (pooled) and fellow: - Acceleration range: .12 ($P = .4$) - Mean jerk: .11 ($P = .5$)

(continued)

Table 3 *(continued)*

Study	Instrument used	Intrarater/test–retest	Interrater
Vaidya et al[33]	Smartphone accelerometer during PS	ICC (95% CI): - AP, intact: .67 (.24 to .87, $P = .005$) - AP, ACLD: .85 (.66 to .94, $P < .001$) - SI, intact: .63 (.19 to .85, $P = .007$) - SI, ACLD: .93 (.85 to .97, $P < .001$) - ML, intact: −.05 (−1.15 to .57, $P =$ n.s.) - ML, ACLD: −.15 (−1.64 to .56, $P =$ n.s.)	ICC (95% CI): - AP, intact: .52 (−.32 to .83, $P =$ n.s.) - AP, ACLD: .95 (.85 to .98, $P < .001$) - SI, intact: .63 (.002 to .87, $P = .027$) - SI, ACLD: .83 (.54 to .94, $P = .001$) - ML, intact: −.01 (−2.06 to .65, $P =$ n.s.) - ML, ACLD: −.09 (−2.30 to .62, $P =$ n.s.)
Krause et al[44]	Handheld inclinometer during dial test	ICC(3,1) (95% CI): - 30°: .83 (.64 to .92) - 90°: .87 (.73 to .94)	ICC(2,1) (95% CI): - 30°: .74 (.61 to .86) - 90°: .83 (.74 to .91)
Alam et al[34]	RMD with inclinometer during dial test	ICC (95% CI) of difference between readings, for combined readings at 30° and 90° flexion, at 4, 6, and 8 N·m torque: .90 (.8 to 1.0)	
McQuade et al[35]	Genucom knee analysis system	20° flexion, at ±8.13 N·m torque: - CoV: 9% - SD (95% CI): 5.5° (3.87° to 9.69°) 80° flexion, at ±8.13 N·m torque: - CoV: 6% - SD (95% CI): 3.1° (2.12° to 5.35°)	
Wroble et al[49]	Genucom knee analysis system		Results differed significantly between examiners
Mills and Hull[36]	Rotatometer	The test–retest mean difference, at 20 N·m torque: <1.0° Cycle-cycle mean difference, at 20 N·m torque: <0.40°	
Chung et al[37]	Rotatometer	ICC calculated 3 times at 1-mo intervals: - IR: .87–.95 ($P < .05$) - ER: .69–.89 ($P < .05$) Highest values in the third month	ICC: - IR: .90–.95 ($P < .05$) - ER: .84–.91 ($P < .05$)
Lorbach et al[38]	Rotameter	ICC: - At 5 N·m torque: IR: .67; ER: .81; IR + ER: .92 - At 10 N·m torque: IR: .79; ER: .88; IR + ER: .94 - At 15 N·m torque: IR: .83; ER: .93; IR + ER: .84	ICC: - At 5 N·m torque: IR: .94; ER: .94; IR + ER: .96 - At 10 N·m torque: IR: .97; ER: .95; IR + ER: .97 - At 15 N·m torque: IR: .98; ER: .95; IR + ER: .88
Neumann et al[39]	Rotameter, updated prototype	The difference in end-range IR, at 15 N·m torque, measured 4 mo apart: 1.6° The difference in end-range ER, at 15 N·m torque, measured 4 mo apart: 5.3° No difference could be seen between the same-day measurements	The difference in end-range IR: 1° The difference in end-range ER: 1.4°
Lorenz et al[40]	KUKA robot	ICC(3,*k*), at 4 N·m torque: .89 ICC(3,1), at 4 N·m torque: .94 All measurement conditions CoV < 5%	
Almquist et al[41]	Rottometer	ICC(2,1) (95% CI), at 3 different knee angles and at 3, 6, and 9 N·m torque, respectively, within a day: - 90°: .64 (.13 to .86), .73 (.43 to .87), .79 (.54 to .91) - 60°: .67 (.33 to .85), .76 (.50 to .90), .86 (.62 to .95) - 30°: .59 (.21 to .81), .87 (.72 to .95), .94 (.87 to .98) 1 wk apart: - 90°: .49 (.04 to .77), .67 (.31 to .86), .80 (.56 to .92) - 60°: .24 (−.15 to .59), .50 (.11 to .76), .82 (.60 to .92) - 30°: .22 (−.49 to .40), .50 (.11 to .77), .84 (.65 to .93)	ICC(2,1) (95% CI), at 3 different knee angles and at 3, 6, and 9 N·m torque, respectively: - 90°: .49 (.09 to .76), .87 (.71 to .95), .85 (.68 to .94) - 60°: .75 (.47 to .89), .83 (.62 to .93), .84 (.63 top .93) - 30°: .52 (.10 to .78), .61 (.25 to .82), .69 (.34 to .87)
Nascimento et al[42]	DYNEELAX	ICC(2,1) (95% CI)/SEM: - IR at 5 N·m: .94 (.91 to .96)/1.08 - ER at 5 N·m: .96 (.96 to .97)/1.04	
Kang et al[43]	Computational modeling derived from Telos stress test	ICC: .99	ICC: .98

Abbreviations: ACLD, anterior cruciate ligament deficient; AP, anterior–posterior; CoV, coefficient of variation; CI, confidence interval; ER, external rotation; IR, internal rotation; ICC, intraclass correlation coefficient; KiRA, Kinematic Rapid Assessment; MRI, magnetic resonance imaging; ML, mediolateral; NR, not reported; n.s., not significant; PS, pivot shift; RKT, robotic knee testing; RMS, root mean square; RMD, rotational measurement device; SEM, standard error measurement; SI, superior–inferior; VKLD, Vermont knee laxity device.

[a]Compliance was defined as the instantaneous slope of the curve at the position of zero torque

Reliability

Intrarater reliability was reported in 25 (59.5%) studies (see Table 3).[20–44] ICCs were reported in 18 (42.9%) studies, and almost all demonstrated adequate reliability (ICC > .75). ICCs that scored below .75 typically involved applying lower torques (eg, 6 N·m or less),[38,41] applying torques while also applying axial compressive loads,[28] or measuring accelerations instead of laxity.[31,33] SEMs were reported along with ICCs in only 5 of the 18 studies reporting ICCs, with values ranging from 0.14°[29] to 3.8°.[30]

Interrater reliability was reported in 15 (35.7%) studies.[20,21,30,33,37–39,41,43–49] ICCs were reported in 12 (28.6%) of the studies, and again, almost all demonstrated adequate reliability (ICC > .75), with the exceptions involving applying lower torques[57] or measuring accelerations.[33,46,48] SEMs were reported in only one study and ranged from 1.6° to 2.6°.[30]

Validity, Continuous Outcomes

Ten different criterion techniques were used among the 19 (45.2%) validity studies with continuous outcomes, of which the pivot-shift test (grade) was most common (Table 4).[20,25,26,34,36,43,45,46,48,50–59] Only 4 (9.5%) studies found *very good* correlations ($r > .8$) with their respective criterion measures.[34,53,54,57] The highest correlations were between polymer-based capacitive strain gauges compared with mechanical rotation ($r = 1$),[54] the Knee Joint Laxity Testing Device using optical motion capture compared with fluoroscopy ($r = .97$),[53] and the Rottometer with knees at 90° flexion applying 9 N·m of torque compared with Roentgen Stereometric Analysis ($r = .97$).[57] When comparing 2 rotational methods, correlations were lowest in two comparisons: Kinematic Rapid Assessment (KiRA) during pivot shift compared with pivot-shift grade (no correlation for jerk; no correlation for range when performed by orthopedic fellow)[48] and electromagnetic motion tracking of external tibial rotation during the pivot-shift test compared with pivot-shift grading ($r = .16$).[46]

Some studies included anteroposterior stability tests as a comparison measure. The highest correlation was between the measurement of anteromedial rotary displacement of the lateral compartment during Slocum test in an open-bore MRI compared with the KT-2000 ($r = .48$).[51] The lowest correlations were between Slocum test in an open-bore MRI and the side-to-side difference on the KT-2000 ($r = .08$–$.13$).[50]

Six (14.9%) studies reported known groups validity.[25,48,50,52,56,59] All studies showed a significant difference between groups with intact versus injured knees, with intact knees having the best rotational stability, followed by ACL-reconstructed knees, and the least stability was in anterior cruciate ligament-deficient (ACLD) knees. A study evaluating the KiRA in injured compared with contralateral uninjured knees only detected a significant difference between groups when patients were under anesthesia but not when patients were awake.[59]

Validity, Dichotomous Outcomes

Nine (21.4%) studies evaluated dichotomous outcomes, 6 (14.3%) of which evaluated both sensitivity and specificity (Table 5).[32,33,45,48,50–52,60,61] The highest sensitivity was 100% for Slocum test in an open-bore MRI.[50] The lowest sensitivity was 25% for the Rotameter, which required a threshold for both internal rotation and compliance, although this test also had the highest specificity of 100%.[61] The Smartphone accelerometer also had the highest specificity of 100% in the superoinferior direction but also the lowest specificity of 35% in the mediolateral direction.[33] Three (7.1%) studies reported AUC[33,51,60] and reported acceptable AUCs (>0.7), with the exception of the Smartphone accelerometer in mediolateral and anteroposterior directions.[33] Agreement was reported in 4 (9.5%) studies,[45,48,51,61] with the highest agreement of 84.8% for Slocum test with MRI compared with the pivot-shift test.[51]

Risk of Bias Assessments

Only one of the 30 reliability studies scored *very good* on risk of bias (Table B1 in Appendix B).[36] Four (9.5%) studies scored *adequate,*[20,31,46,47] 16 (38.1%) scored *doubtful,*[21,22,24,27–30,33,34,38,40–44,48] and 9 (21.4%) *inadequate.*[23,25,26,32,35,36,39,45,49] All studies scored *adequate or good* in the "design requirements" domain. Ten (23.8%) studies scored *doubtful* or *inadequate* in the "statistical methods" domain, mostly because studies did not report an ICC or Pearson or Spearman correlation. Twenty-two (52.4%) studies scored *doubtful* in the "other" domain, typically due to small sample sizes.

Six (14.2%) validity studies with continuous outcomes scored *very good* (Table B2 in Appendix B).[20,46,48,51,52,56] None scored *adequate*, and 13 (31.0%) scored *doubtful*[25,26,34,50,53,57–59] or *inadequate*.[36,43,45,54,55] Four (9.5%) studies scored *inadequate* in the "statistical methods" domain because they did not report a correlation or AUC.[36,43,45,55] Among the validity studies with dichotomous outcomes, only 1 was rated as having *low* risk of bias,[52] whereas 5 (11.9%) were rated as *high* risk of bias,[32,33,48,60,61] and 3 (7.1%) were rated as unclear (Table B3 in Appendix B).[45,50,51] Risk of bias was mainly due to potential prior knowledge of the knee condition when performing the index tests.

Discussion

Measurement instruments for knee rotational stability varied from small handheld devices to extensive apparatuses requiring different levels of human intervention. Instruments used were generally of adequate reliability (ICCs > 0.75), particularly when higher torques of greater than 6 N·m were applied to the knee. Four (9.5%) studies reported *very good* correlations ($r > .8$) for 4 different instruments compared with 4 different criterion measures.[34,53,54,57] Both reliability and validity were assessed in ten different devices/setups. Of these, 3 studies evaluating rotational measurement devices were reported to have both adequate reliability and very good validity, as was 1 study of a setup that used computational modeling based on a Telos stress test.[34,41,43,57] However, risk of bias in these studies was rated as "doubtful" in 3 (7.1%) studies and "inadequate" in 1 (2.4%) study. The remaining 6 (14.3%) studies that did not demonstrate sufficient reliability and/or validity all evaluated devices that involved the performance of a clinical test (ie, Slocum or pivot shift). No high-quality evidence provided sufficient and comprehensive evidence of both reliability (eg, both ICC and SEM reported) and validity (eg, both correlations and mean error compared with an acceptable criterion standard, eg, fluoroscopy).

Study heterogeneity made direct comparison among devices difficult. Device design along with several other factors may explain differing results among studies. For example, reliability was generally higher when evaluating individuals with ACLD compared with intact knees.[31,33] Reliability was also higher with higher applied torques. Devices that measured movements while a tester performed a pivot-shift test did not perform as well as devices

Table 4 Validity

Study	Instrument used	Criterion method (condition)	Validity	*P*
Okazaki et al[20]	Horizontally open-bore MRI during Slocum test:			
	Anteromedial rotary displacement of the lateral compartment	Anterior displacement during stress radiography using Telos stress test	ACLD: $r = .67$	.005
Tashiro et al[50]	Horizontally open-bore MRI during Slocum test:			
	Anteromedial rotary displacement of lateral compartment, ACL vs intact contralateral side difference	PS grade	ACLR: $r = .74$ ACLD: $r = .55$	<.01 <.05
		KT-2000 (side-to-side difference)	ACLR: $r = .13$ ACLD: $r = .08$	.59 .75
		Stress radiography using Telos (side-to-side difference)	ACLR: $r = .54$ ACLD: $r = .46$	<.05 <.05
		Known groups validity	Intact: 1.0 (2.3) mm ACLR: 2.1 (3.8) mm	<.01
			Intact: 1.6 (3.2) mm ACLD: 9.2 (5.5) mm	<.01
Okazaki et al[21]	Horizontally open-bore MRI during Slocum test:			
	Anteromedial rotary displacement of the lateral compartment	PS grade KT-2000	$r = .75$ $r = .48$	<.001 <.001
	Difference between translation applied to lateral tibiofemoral vs medial tibiofemoral compartment	PS grade KT-2000	$r = .65$ $r = .36$	<.001 .0014
Espregueira-Mendes et al[52]	MRI with PKTD:			
	LPpa	PS	NR	n.s.
	MPpa	PS	NR	n.s.
	LPpa – LPns	PS	$r = .34$	.005
	MPpa – MPns	PS	NR	n.s.
	LPpa + MPpa	PS	NR	n.s.
	LPpair + LPpaer	PS	$r = .55$	.001
		Known groups validity:		
		MPpair	Intact: –0.9 (3.8) mm ACLD: 6.3 (3.5) mm	<.001 <.001
		LPpair	Intact: 6.24 (4.1) mm ACLD: 15.6 (5.6) mm	<.001 <.001
		MPpaer	Intact: 1.6 (2.7) mm ACLD: 9.5 (5.0) mm	<.001 <.001
		LPpaer	Intact: 9.5 (5.0) mm ACLD: 9.5 (5.0) mm	<.001 <.001
		LPpair + LPpaer	Intact: 9.7 (8.9) mm ACLD: 28.0 (10.5) mm	<.001 <.001
		MPpa + LPpair	Intact: 28.0 (10.5) mm ACLD: 28.0 (10.5) mm	<.001 <.001
Hoshino et al[25]	Electromagnetic measurement during PS:			
	Coupled anterior translation	PS	Intact: NR ACLR: NR	.03 .14
		KT-1000	Intact: NR ACLR: $r = .26$	.17
	Acceleration of posterior translation	PS	Intact: NR ACLR: NR	<.01 .14

(continued)

Table 4 *(continued)*

Study	Instrument used	Criterion method (condition)	Validity	*P*
		KT-1000	Intact: NR ACLR: NR	
		Known groups validity:		
		Coupled anterior translation	Intact: 7.7 (0.6) mm ACLR: 15.6 (1.2) mm	<.01
		Acceleration of posterior translation	Intact: −797 (45) mm/s^2 ACLR: −2001 (186) mm/s^2	<.01
Kubo et al[26]	Electromagnetic measurement during PS:			
	Posterior translation	PS grade	Spearman $r = .28$	<.001
	Lateral translation	PS grade	Spearman $r = .28$	<.001
	Maximum velocity	PS grade	Spearman $r = .39$	<.001
Labbe et al[46]	Electromagnetic measurement during PS:			
	Tibial ER	PS grade	Spearman $r = .16$	.01
	Acceleration of tibial rotation	PS grade	Spearman $r = .21$	.01
	Posterior translation	PS grade	Spearman $r = .22$	.01
	Acceleration of posterior translation	PS grade	Spearman $r = .42$	.01
Labbe et al[45]	Support vector machine algorithm based on electromagnetic measurement during PS (grades)	PS grade	$K = 0.68$	
Moewis et al[53]	Optical motion capture system with knee joint laxity testing device	Knee joint laxity testing device with fluoroscopy	Full extension: $r = .97$ Mean rotations, full extension: motion capture 24.2 (2.6)° vs fluoroscopy 9.9 (2.1)° Bland–Altman LoA without/with correction: 0.2° (−20.8°, 20.3°)/0.0° (−0.7°, 0.7°) RMSE without/with correction: 14.2°/2.3° (full extension had largest error, gradually reduced with knee flexion)	
Mayr et al[54]	Polymer-based sensor elastic–capacitive measurement system IR measured at the foot	Laxitester	Spearman $r = 1$	<.01
Bellitti et al[55]	Smart brace with IMUs, during PS	Marker-based stereophotogrammetric system (optoelectronic)	Absolute angle difference between the 2 methods: 13° Absolute average difference between the 2 methods: 21°	
Nakamura et al[59]	KiRA during PS:			
	Acceleration range	PS grade	Anesthesia: $r = .57$ Awake: $r = .41$	.0014 .0027
		Known groups validity:	Injured, preoperative vs uninjured contralateral:	
		Anesthesia	NR	<.001
		Awake	NR	n.s.
Helfer et al[56]	KiRA during PS:			
	Acceleration range	PS grade	$r = .57$ (95% CI = .51 to .66)	<.001
		Known groups validity:	Grade 2 vs contralateral: 3.5 (1.6) m/s^2	<.001
		Side-to-side difference	Grade 3 vs contralateral: 5.9 (3.6) m/s^2	<.001
Napier et al[48]	KiRA during PS:			
	Acceleration range	PS grade	Surgeon: $r = .40$ Fellow: no correlation	<.01
	Mean jerk	PS grade	No correlation for both surgeon and fellow	<.01
		Known groups validity for the acceleration range	Grade 1: Surgeon: 3.91 (1.49) m/s^2	.01

(continued)

Table 4 *(continued)*

Study	Instrument used	Criterion method (condition)	Validity	*P*
			Fellow: 5.08 (0.51) m/s^2	
			Grade 2:	.01
			Surgeon: 4.24 (1.64) m/s^2	
			Fellow: 5.01 (1.63) m/s^2	
			Grade 3:	.01
			Surgeon: 6.38 (0.84) m/s^2	
			Fellow: 5.52 (2.75) m/s^2	
Alam et al[34]	RMD with inclinometer during dial test:			
	Rotation measured at the tibia	Nest-of-birds electromagnetic tracking system	r = .92 (95% CI = .89 to .94) Mean difference: –2° (95% CI = 1°to –4°)	
	Rotation measured at the foot	Nest-of-birds electromagnetic tracking system	r = .63 (95% CI = .54 to .70) Mean difference: –34° (95% CI–9°to –58°)	
Mills and Hull[36]	Rotatometer: IR/ER	Custom goniometer	Mean absolute error: 0.5°	
Almquist et al[57]	Rottometer:	RSA		
	IR + ER		3 N·m: r = .93	<.001
			6 N·m: r = .96	<.001
			9 N·m: r = .97	<.001
			90°: r = .97	<.001
			60°: r = .96	<.001
	IR		3 N·m: r = .93	<.001
			6 N·m: r = .69	<.03
			9 N·m: r = .74	<.014
			90°: r = .93	<.001
			60°: r = .76	<.001
	ER		3 N·m: r = .69	<.028
			6 N·m: r = .88	<.001
			9 N·m: r = .71	<.02
			90°: r = .73	<.001
			60°: r = .81	<.001
Mayr et al[58]	Laxitester with IR and ER applied at the foot	PS grade	Primary ACL tear: r = .695	<.001
			ACL rerupture: r = .637	<.001
Kang et al[43]	Computational modeling derived from Telos stress test:		The mean absolute difference among 22 chosen points:	
	Initial material properties model	CT	10.39 (8.26 mm)	
	Optimized material properties model	CT	4.21 (2.56) mm	

Abbreviations: ACL, anterior cruciate ligament; ACLD, anterior cruciate ligament deficient; ACLR, anterior cruciate ligament reconstructed; CI, confidence interval; CT, computed tomography; ER, external rotation; IR, internal rotation; KiRA, Kinematic Rapid Assessment; LPns, lateral tibia plateau no stress; LPpa, lateral tibial plateau with posteroanterior translation (stress applied to whole knee); LPpair, lateral tibial plateau with posteroanterior translation coupled with maximum internal rotation; LPpaer, lateral tibial plateau with posteroanterior translation coupled with maximum external rotation; LoA, limits of agreement; MRI, magnetic resonance imaging; MPns, medial tibia plateau no stress; MPpaer, medial tibial plateau with posteroanterior translation coupled with maximum external rotation; MPpa, medial tibial plateau with posteroanterior translation; NR, not reported; n.s., not significant; PS, pivot shift; PKTD, Porto knee testing device; RSA, Roentgen stereometric analysis; RMSE, root mean square error; RMD, rotational measurement device.

that also applied a standardized torque, though setups involving pivot shift were sensitive to tester experience (more experience improved reliability)[32,48] and speed of the pivot-shift maneuver (higher speed was more reliable).[26,31,46] Devices that measured the amount of rotation in response to applied torques also tended to be more reliable than measures of acceleration or jerk. This may be because these devices measure a pivot-shift test, which is sensitive to tester experience and maneuver speed, as described earlier. It may also be because these are primarily sensor-based devices, and their placement on the skin may introduce soft-tissue artifact, or the algorithms need further calibration and refinement.

It would be expected that without anesthesia, individuals may resist unnatural or uncomfortable knee movements with muscle guarding, which may affect rotational stability testing. However, studies that tested rotation with participants under anesthesia did not demonstrate substantially higher reliability or validity than studies in which no anesthesia was given. Only one (2.4%) study directly assessed correlations between the KiRA and pivot-shift grading with and without anesthesia, and although correlations were higher under anesthesia, they remained fair under both conditions.[59] It is, thus, unclear whether testing under anesthesia while evaluating rotational knee stability adds clinical value.

Three (7.1%) studies reported poor to fair correlations between rotational stability and translational stability, as measured with the KT-1000 or KT-2000.[25,50,51] Lower correlations between rotational and translational stability would, indeed, be expected in

Table 5 Diagnostic Accuracy

Study	Instrument used	Criterion method	Agreement	Sensitivity/specificity	AUC (*P*)	Cutoff
Tashiro et al[50]	Horizontally open-bore MRI during Slocum test	ACLD vs ACLR		100%/92%		3 mm side-to-side difference
Okazaki et al[51]	Horizontally open-bore MRI during Slocum test (PA translation of lateral knee)	Clinical grading during PS of ACLD vs ACLR knees	84.8% 70.7%		PS grade 1: 0.98 PS grade 2: 0.95	PS grade 1: 3.2 mm PS grade 2: 6.4 mm
Espregueira-Mendes et al[52]	MRI with PKTD	Clinical signs and symptoms of unilateral ACLD knee, referred for MRI vs contralateral intact knee				
	(PA-translation medial plateau)			72.1%/87.5%		<3.5 mm
	(PA-translation lateral plateau)			73.8%/81.3%		<4.3 mm
	(Global translation: PA translation medial plus lateral plateau)			86.9%/93.8%		<11.1 mm
	(Global rotation, lateral plateau: PA translation with maximum internal rotation plus PA translation with maximum external rotation)			92.9%/80.0%		<15.1 mm
	(Global rotation medial plateau: PA translation with maximum internal rotation plus PA translation with maximal external rotation)			88.5%/93.8%		<12.1 mm
Araki et al[60]	Electromagnetic measurement during PS	Arthroscopic inspection to grade fully ACLD, partial ACLD, and intact knees		Partial ACLD: 85%/80% ACLD: 75%/75%	Partial ACLD: 0.88 ACLD: 0.75	Partial ACLD: −848.8 mm/s^2 ACLD: 1245.3 mm/s^2
Labbe et al[45]	Support vector machine algorithm based on electromagnetic measurement during PS	PS grading by an orthopedic surgeon for fully and partial ACLD knees and intact ACL	Weighted $\kappa = .68$	86%/90%		
Berruto et al[32]	KiRA during PS	Doctor diagnosis		Specificity improved linearly from 50% to 90% over a 9-mo study (learning curve)		
Napier et al[48]	KiRA during PS (Acceleration range) (Mean jerk)	The operating surgeon and orthopedic fellow agree on which knee is ACLD (patients with unilateral ACL tear)	Surgeon or fellow: 74% Surgeon: 76% Fellow: 80%			
Vaidya et al[33]	Smartphone accelerometer during PS (Acceleration range)	Unilateral ACLD knee awaiting ACL reconstruction (vs intact knee)		AP: 53%/65% SI: 94%/100% ML: 29%/35%	AP: 0.54 SI: 0.98 ML: 0.25	AP: −0.05 m/s^2 SI: 1.24 m/s^2 ML: −0.003 m/s^2

(continued)

Table 5 ***(continued)***

Study	Instrument used	Criterion method	Agreement	Sensitivity/specificity	AUC (*P*)	Cutoff
Mouton et al[61]	Rotameter	Arthroscopically confirmed unilateral ACLD knee awaiting reconstruction	IR with 5 N·m torque: 63% accuracy	IR with 5 N·m torque: 38%/95%		IR: ≥3.2°
			Compliance in IR from 2 to 5 N·m torque: 59% accuracy	Compliance in IR from 2 to 5 N·m torque: 31%/95%		Compliance in IR: 0.6°/N·m
			Both (at least 1 above threshold): 64% accuracy	Both (at least 1 above threshold): 44%/90%		At least 1 above threshold
			Both (2 above threshold): 58% accuracy	Both (2 above threshold): 25%/100%		Both above threshold

Abbreviations: AUC, area under the receiver operating curve; ACLD, anterior cruciate ligament deficient; ACLR, anterior cruciate ligament reconstructed; ACL, anterior cruciate ligament; AP, anterior–posterior; IR, internal rotation; KiRA, Kinematic Rapid Assessment,; MRI, magnetic resonance imaging; ML, medial–lateral; PS, pivot shift; PKTD, Porto knee testing device; PA, posteroanterior; SI, superior–inferior.

comparison with correlations between 2 measures of rotational stability. These results, thus, provide additional evidence of construct validity of these instruments.

The outcome parameters reported and criterion methods among studies differed given that a gold standard does not exist for this measurement. This does make comparisons among studies difficult, particularly as the criterion test was often a manually performed clinical test, calling into question which approach was actually more valid (instrument or clinical test) in cases of lower correlations. Although ICCs were the most commonly reported reliability statistic, studies rarely reported SEMs, which should always be reported alongside ICCs to enhance interpretability of these statistics.

Our study synthesizes and summarizes the reliability and validity of devices designed to measure knee stability in a rotational direction. This adds to existing knowledge, which has focused on instruments designed to evaluate knee stability in an anteroposterior direction. A previous systematic review concluded that the KT-1000 and KT-2000 were the most accurate and reliable instruments for evaluating knee stability in the anteroposterior direction.[7] Based on these new findings, 2 different instruments would be required to objectively measure stability in both planes. This could pose feasibility issues in a clinical setting in that clinics would need to purchase 2 devices, and extra time would be required during a clinic visit to set up and use the 2 devices to take multiple measurements. Based on the available literature, we cannot yet confirm the clinical benefit of implementing such an objective rotational assessment into a standard clinical evaluation. Still, given the complex multiplanar nature of knee instability in the case of ACLD, we hypothesize that more objectively measuring stability in multiple planes would meaningfully support clinical decision making. To evaluate the clinical benefit of implementing objective multiplanar assessment into clinical care, it is first necessary to identify or design an instrument with sufficient validity and reliability. Our present review suggests that this has not yet been achieved.

Limitations and Future Research

First, we did not include gray literature in this review; therefore, we may have missed relevant studies that had not yet undergone peer review. Second, the pivot-shift test was the most common criterion method, which relies heavily on a subjective examiner.[6] This reduces the credibility of the relevant included validation studies, although we recognize the high specificity of this test for ACLD.[4] Finally, most instruments were only tested in 1 study and often by the developers themselves. Future independent research should focus on systematically testing both reliability and validity in adequately powered studies, particularly for the instruments showing promising results thus far, and in adherence with best practices.[14,17,18] Reliability assessments should report SEM along with ICC as the ICC statistic is unitless. The SEM provides a relative unit of measure and is, thus, critical to aid in interpreting the amount of error relative to a given range of possible values that can be measured.[14] Validation should include both construct and concurrent validation compared with an acceptable criterion standard.[14] Not all instruments evaluated in the present review are commercially available, making them challenging to obtain for future research or clinical implementation.

Conclusions

Instruments designed to measure rotational knee stability are typically designed either to objectively measure movement while applying a standardized rotational force or to objectively measure movement while performing manual testing, such as a pivot-shift test. Among 42 studies that met our eligibility criteria, we found no high-quality evidence that provided sufficient and comprehensive evidence of both reliability and validity in any device. Devices measuring rotational knee laxity demonstrated adequate reliability, particularly when torques of greater than 6 N·m were applied to the knee. Devices measuring other aspects of rotational stability, like acceleration or jerk, were less reliable. Only 4 devices demonstrated sufficient concurrent validation, but all 4 studies were of doubtful or inadequate risk of bias. Future studies should evaluate reliability and validity together, use a criterion measure such as fluoroscopy against which to evaluate validity, and adopt best practice reporting; for example, reporting SEM along with ICCs, which is critical to aid in clinical interpretation.[18] The aim of this research should be to inform the design of a valid and reliable instrument that is also feasible to use in clinical settings. This will make it possible to evaluate the clinical benefit of objectively evaluating rotational stability alongside anteroposterior stability with an ultimate goal of improving clinical assessment and decision making.

References

1. Griffin LY, Agel J, Albohm MJ, et al. Noncontact anterior cruciate ligament injuries: risk factors and prevention strategies. *J Am Acad Orthop Surg*. 2000;8(3):141–150. PubMed ID: 10874221 doi:10.5435/00124635-200005000-00001
2. Noyes FR. The function of the human anterior cruciate ligament and analysis of single- and double-bundle graft reconstructions. *Sports Health*. 2009;1(1):66–75. PubMed ID: 23015856 doi:10.1177/1941738108326980
3. Huang Z, Liu Z, Fan C, Zou M, Chen J. Value of clinical tests in diagnosing anterior cruciate ligament injuries: a systematic review and meta-analysis. *Medicine*. 2022;101(31):e29263 . doi:10.1097/MD.0000000000029263
4. Prins M. The Lachman test is the most sensitive and the pivot shift the most specific test for the diagnosis of ACL rupture. *Aust J Physiother*. 2006;52(1):66 . doi:10.1016/S0004-9514(06)70069-1
5. Slichter ME, Wolterbeek N, Auw Yang KG, Zijl JAC, Piscaer TM. Rater agreement reliability of the dial test in the ACL-deficient knee. *J Exp Orthop*. 2018;5:18 . doi:10.1186/s40634-018-0131-y
6. Lane CG, Warren R, Pearle AD. The pivot shift. *J Am Acad Orthop Surg*. 2008;16(12):679–688. PubMed ID: 19056917 doi:10.5435/00124635-200812000-00001
7. Pugh L, Mascarenhas R, Arneja S, Chin PYK, Leith JM. Current concepts in instrumented knee-laxity testing. *Am J Sports Med*. 2009;37(1):199–210. PubMed ID: 18940931 doi:10.1177/0363546508323746
8. Abulhasan JF, Snow MD, Anley CM, Bakhsh MM, Grey MJ. An extensive evaluation of different knee stability assessment measures: a systematic review. *J Funct Morphol Kinesiol*. 2016;1(2):209–229. doi:10.3390/jfmk1020209
9. Ahldén M, Hoshino Y, Samuelsson K, Araujo P, Musahl V, Karlsson J. Dynamic knee laxity measurement devices. *Knee Surg Sports Traumatol Arthrosc*. 2012;20(4):621–632. PubMed ID: 22210516 doi:10.1007/s00167-011-1848-7
10. Lorbach O, Brockmeyer M, Kieb M, Zerbe T, Pape D, Seil R. Objective measurement devices to assess static rotational knee laxity: focus on the Rotameter. *Knee Surg Sports Traumatol Arthrosc*. 2012;20(4):639–644. PubMed ID: 22246547 doi:10.1007/s00167-011-1876-3

11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71 . doi:10.1136/bmj.n71
12. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240 . doi:10.5195/jmla.2016.24
13. The Endnote Team. *Endnote 21*. Clarivate Analytics; 2023.
14. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737–745. doi:10.1016/j.jclinepi.2010.02.006
15. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 4th ed. Oxford University Press; 2008.
16. Chan YH. Biostatistics 104: correlational analysis. *Singapore Med J*. 2003;44(12):614–619. PubMed ID: 14770254
17. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315–1316. PubMed ID: 20736804 doi:10.1097/JTO.0b013e3181ec173d
18. Mokkink LB, De Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1171–1179. PubMed ID: 29260445 doi:10.1007/s11136-017-1765-4
19. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–536. PubMed ID: 22007046 doi:10.7326/0003-4819-155-8-201110180-00009
20. Okazaki K, Miura H, Matsuda S, et al. Assessment of anterolateral rotatory instability in the anterior cruciate ligament—deficient knee using an open magnetic resonance imaging system. *Am J Sports Med*. 2007;35(7):1091–1097. PubMed ID: 17379920 doi:10.1177/0363546507299530
21. Tardy N, Marchand P, Kouyoumdjian P, Blin D, Demattei C, Asencio G. A preliminary in vivo assessment of anterior cruciate ligament–deficient knee kinematics with the KneeM device: a new method to assess rotatory laxity using open MRI. *Orthop J Sports Med*. 2014;2(3). doi:10.1177/2325967114525583
22. Hemmerich A, van der Merwe W, Vaughan CL. Measuring three-dimensional knee kinematics under torsional loading. *J Biomech*. 2009;42(2):183–186. PubMed ID: 19059595 doi:10.1016/j.jbiomech.2008.10.021
23. Carpenter RD, Shefelbine SJ, Lozano J, Carballido-Gamio J, Majumdar S, Benjamin Ma C. A new device for measuring knee rotational kinematics using magnetic resonance imaging. *J Med Devices*. 2008;2(4):044501 . doi:10.1115/1.2976029
24. Haughom BD, Souza R, Schairer WW, Li X, Benjamin Ma C. Evaluating rotational kinematics of the knee in ACL-ruptured and healthy patients using 3.0 Tesla magnetic resonance imaging. *Knee Surg Sports Traumatol Arthrosc*. 2012;20(4):663–670. PubMed ID: 22159520 doi:10.1007/s00167-011-1809-1
25. Hoshino Y, Kuroda R, Nagamune K, et al. In vivo measurement of the pivot-shift test in the anterior cruciate ligament—deficient knee using an electromagnetic device. *Am J Sports Med*. 2007;35(7):1098–1104. PubMed ID: 17351123 doi:10.1177/0363546507299447
26. Kubo S, Muratsu H, Yoshiya S, Mizuno K, Kurosaka M. Reliability and usefulness of a new in vivo measurement system of the pivot shift. *Clin Orthop Relat Res*. 2007;454:54–58. PubMed ID: 17091016 doi:10.1097/BLO.0b013e31802b4a38
27. Kuroda R, Hoshino Y, Kubo S, et al. Similarities and differences of diagnostic manual tests for anterior cruciate ligament insufficiency: a global survey and kinematics assessment. *Am J Sports Med*. 2012;40(1):91–99. PubMed ID: 21989128 doi:10.1177/0363546511423634
28. Shultz SJ, Shimokochi Y, Nguyen AD, Schmitz RJ, Beynnon BD, Perrin DH. Measurement of varus–valgus and internal–external rotational knee laxities in vivo—part I: assessment of measurement reliability and bilateral asymmetry. *J Orthop Res*. 2007;25(8):981–988. PubMed ID: 17457828 doi:10.1002/jor.20397
29. Beckley S, Stinton S, Lesosky M, et al. Reliability of a robotic knee testing tool to assess rotational stability of the knee joint in healthy female and male volunteers. *Sports Med—Open*. 2020;6(1):33 . doi:10.1186/s40798-020-00266-7
30. Tsai AG, Musahl V, Steckel H, et al. Rotational knee laxity: reliability of a simple measurement device in vivo. *BMC Musculoskelet Disord*. 2008;9:35 . doi:10.1186/1471-2474-9-35
31. Lopomo N, Zaffagnini S, Signorelli C, et al. An original clinical methodology for non-invasive assessment of pivot-shift test. *Comput Methods Biomech Biomed Engin*. 2012;15(12):1323–1328. PubMed ID: 21728739 doi:10.1080/10255842.2011.591788
32. Berruto M, Uboldi F, Gala L, Marelli B, Albisetti W. Is triaxial accelerometer reliable in the evaluation and grading of knee pivot-shift phenomenon? *Knee Surg Sports Traumatol Arthrosc*. 2013;21(4):981–985. PubMed ID: 23392289 doi:10.1007/s00167-013-2436-9
33. Vaidya RK, Yoo CW, Lee J, Han H-S, Lee MC, Ro DH. Quantitative assessment of the pivot shift test with smartphone accelerometer. *Knee Surg Sports Traumatol Arthrosc*. 2020;28(8):2494–2501. PubMed ID: 31840191 doi:10.1007/s00167-019-05826-3
34. Alam M, Bull AMJ, Thomas Rd, Amis AA. A clinical device for measuring internal–external rotational laxity of the knee. *Am J Sports Med*. 2013;41(1):87–94. PubMed ID: 23277467 doi:10.1177/0363546512469874
35. McQuade KJ, Sidles JA, Larson RV. Reliability of the genucom knee analysis system: a pilot study. *Clin Orthop Relat Res*. 1989;245:216–219. doi:10.1097/00003086-198908000-00033
36. Mills OS, Hull ML. Apparatus to obtain rotational flexibility of the human knee under moment loads in vivo. *J Biomech*. 1991;24(6):351–369. PubMed ID: 1856237 doi:10.1016/0021-9290(91)90025-I
37. Chung JH, Ryu KJ, Lee DH, et al. An analysis of normative data on the knee rotatory profile and the usefulness of the Rotatometer, a new instrument for measuring tibiofemoral rotation: the reliability of the knee Rotatometer. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(9):2727–2733. PubMed ID: 24817107 doi:10.1007/s00167-014-3039-9
38. Lorbach O, Wilmes P, Theisen D, et al. Reliability testing of a new device to measure tibial rotation. *Knee Surg Sports Traumatol Arthrosc*. 2009;17(8):920–926. PubMed ID: 19305971 doi:10.1007/s00167-009-0772-6
39. Neumann S, Maas S, Waldmann D, et al. Design, repeatability, and comparison to literature data of a new noninvasive device called "rotameter" to measure rotational knee laxity. *Int Sch Res Notices*. 2015;2015(1):439095 . doi:10.1155/2015/439095
40. Lorenz A, Krickl V, Ipach I, Arlt E-M, Wülker N, Leichtle UG. Practicability for robot-aided measurement of knee stability in-vivo. *BMC Musculoskelet Disord*. 2015;16(1):373 . doi:10.1186/s12891-015-0826-5
41. Almquist PO, Ekdahl C, Isberg P-E, Fridén T. Measurements of knee rotation-reliability of an external device in vivo. *BMC Musculoskelet Disord*. 2011;12(1):291 . doi:10.1186/1471-2474-12-291
42. Nascimento N, Kotsifaki R, Papakostas E, et al. DYNEELAX robotic arthrometer reliability and feasibility on healthy and anterior cruciate ligament injured/reconstructed persons. *Transl Sports Med*. 2024;2024:3413466 . doi:10.1155/2024/3413466

43. Kang K-T, Kim S-H, Son J, Lee YH, Chun H-J. Computational model-based probabilistic analysis of in vivo material properties for ligament stiffness using the laxity test and computed tomography. *J Mater Sci Mater Med*. 2016;27(12):183 . doi:10.1007/s10856-016-5797-z
44. Krause DA, Levy BA, Shah JP, Stuart MJ, Hollman JH, Dahm DL. Reliability of the dial test using a handheld inclinometer. *Knee Surg Sports Traumatol Arthrosc*. 2013;21(5):1011–1016. PubMed ID: 22246546 doi:10.1007/s00167-011-1873-6
45. Labbe DR, De Guise JA, Mezghani N, et al. Objective grading of the pivot shift phenomenon using a support vector machine approach. *J Biomech*. 2011;44(1):1–5. PubMed ID: 20810115s doi:10.1016/j.jbiomech.2010.08.012
46. Labbe DR, de Guise JA, Godbout V, et al. Accounting for velocity of the pivot shift test manoeuvre decreases kinematic variability. *Knee*. 2011;18(2):88–93. PubMed ID: 20650637 doi:10.1016/j.knee.2010.03.008
47. Branch TP, Browne JE, Campbell JD, et al. Rotational laxity greater in patients with contralateral anterior cruciate ligament injury than healthy volunteers. *Knee Surg Sports Traumatol Arthrosc*. 2010;18(10):1379–1384. PubMed ID: 20020101 doi:10.1007/s00167-009-1010-y
48. Napier RJ, Feller JA, Devitt BM, et al. Is the KiRA device useful in quantifying the pivot shift in anterior cruciate ligament–deficient knees? *Orthop J Sports Med*. 2021;9(1). doi:10.1177/2325967120977869
49. Wroble RR, Grood ES, Noyes FR, Schmitt DJ. Reproducibility of Genucom knee analysis system testing. *Am J Sports Med*. 1990;18(4):387–395. PubMed ID: 2403188 doi:10.1177/036354659001800410
50. Tashiro Y, Okazaki K, Miura H, et al. Quantitative assessment of rotatory instability after anterior cruciate ligament reconstruction. *Am J Sports Med*. 2009;37(5):909–916. PubMed ID: 19261904 doi:10.1177/0363546508330134
51. Okazaki K, Tashiro Y, Izawa T, Matsuda S, Iwamoto Y. Rotatory laxity evaluation of the knee using modified Slocum's test in open magnetic resonance imaging. *Knee Surg Sports Traumatol Arthrosc*. 2012;20(4):679–685. doi:10.1007/s00167-011-1861-x PubMed ID: PubMed ID: 22210519
52. Espregueira-Mendes J, Andrade R, Leal A, et al. Global rotation has high sensitivity in ACL lesions within stress MRI. *Knee Surg Sports Traumatol Arthrosc*. 2017;25(10):2993–3003. PubMed ID: 27530386 doi:10.1007/s00167-016-4281-0
53. Moewis P, Boeth H, Heller MO, et al. Towards understanding knee joint laxity: errors in non-invasive assessment of joint rotation can be corrected. *Med Eng Phys*. 2014;36(7):889–895. PubMed ID: 24780754 doi:10.1016/j.medengphy.2014.03.017
54. Mayr HO, Rosenstiel N, Prakash KS, et al. Internal rotation measurement of the knee with polymer-based capacitive strain gauges versus mechanical rotation measurement taking gender differences into account: a comparative analysis. *Life*. 2024;14(1):142 . doi:10.3390/life14010142
55. Bellitti P, Borghetti M, Lopomo NF, Sardini E, Serpelloni M. Smart brace for static and dynamic knee laxity measurement. *Sensors*. 2022;22(15):5815 . doi:10.3390/s22155815
56. Helfer L, Vieira TD, Praz C, et al. Triaxial accelerometer evaluation is correlated with IKDC grade of pivot shift. *Knee Surg Sports Traumatol Arthrosc*. 2020;28(2):381–388. PubMed ID: 31201443 doi:10.1007/s00167-019-05563-7
57. Almquist PO, Arnbjörnsson A, Zätterström R, Ryd L, Ekdahl C, Fridén T. Evaluation of an external device measuring knee joint rotation: an in vivo study with simultaneous Roentgen stereometric analysis. *J Orthop Res*. 2002;20(3):427–432. PubMed ID: 12038614 doi:10.1016/S0736-0266(01)00148-6
58. Mayr HO, Hellbruegge G, Haasters F, Ipach B, Schmal H, Prall WC. Laxity measurement of internal knee rotation after primary anterior cruciate ligament rupture versus rerupture. *Arch Orthop Trauma Surg*. 2022;142(10):2839–2847. PubMed ID: 34870728 doi:10.1007/s00402-021-04269-1
59. Nakamura K, Koga H, Sekiya I, et al. Evaluation of pivot shift phenomenon while awake and under anaesthesia by different manoeuvres using triaxial accelerometer. *Knee Surg Sports Traumatol Arthrosc*. 2017;25(8):2377–2383. PubMed ID: 26233597 doi:10.1007/s00167-015-3740-3
60. Araki D, Kuroda R, Matsushita T, et al. Biomechanical analysis of the knee with partial anterior cruciate ligament disruption: quantitative evaluation using an electromagnetic measurement system. *Arthroscopy*. 2013;29(6):1053–1062. PubMed ID: 23602014 doi:10.1016/j.arthro.2013.02.012
61. Mouton C, Theisen D, Meyer T, et al. Combined anterior and rotational knee laxity measurements improve the diagnosis of anterior cruciate ligament injuries. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(10):2859–2867. PubMed ID: 26318487 doi:10.1007/s00167-015-3757-7

Appendix A: Search Terms

The search strings for each of the 4 databases are listed:

PubMed: ("Reproducibility of Results" [MeSH Terms] OR "Sensitivity and Specificity" [MeSH Terms] OR "validation" [Title/Abstract] OR "validity" [Title/Abstract] OR "accuracy" [Title/Abstract] OR "accurately" [Title/Abstract] OR "reliability" [Title/Abstract] OR "reliable" [Title/Abstract] OR "reproducible" [Title/Abstract] OR "reproducibility" [Title/Abstract] OR "repeatability" [Title/Abstract] OR "precision" [Title/Abstract]) AND ("Orthopedic Equipment" [MeSH Major Topic] OR "physical examination/instrumentation" [MeSH Major Topic] OR "arthrometry, articular" [MeSH Terms] OR "arthrometer" [Title/Abstract] OR "arthrometers" [Title/Abstract] OR "device" [Title/Abstract] OR "devices" [Title/Abstract] OR "equipment" [Title/Abstract] OR "laximetry" [Title/Abstract]) AND ("Knee Dislocation" [MeSH Terms] OR "Anterior Cruciate Ligament Injuries" [MeSH Terms] OR "Joint Instability" [MeSH Terms] OR "ACL" [Title/Abstract] OR "anterior cruciate ligament" [Title/Abstract] OR "ligament" [Title/Abstract])

Embase: ('reproducibility'/exp OR' sensitivity and specificity'/exp OR 'validation study'/exp OR 'validation'/exp OR 'validity'/exp OR 'diagnostic accuracy'/exp OR 'reliability'/exp OR 'accuracy'/exp OR 'accurately':ti, ab OR 'reliable':ti,ab OR 'reproducible':ti,ab OR 'precision':ti,ab OR 'repeatability':ti, ab) AND ('orthopedic diagnostic device'/exp OR 'knee arthroscopy'/exp OR 'arthrometry'/exp OR 'arthrometer'/exp OR 'arthrometers':ti,ab OR 'device':ti,ab OR 'devices':ti,ab OR 'equipment':ti,ab OR 'laximetry':ti,ab) AND ('knee dislocation'/exp OR 'knee ligament injury'/exp OR 'knee meniscus injury'/exp OR ('joint stability'/exp AND 'knee') OR 'knee instability'/exp OR ('joint laxity'/exp AND 'knee') OR 'acl':ti, ab OR 'anterior cruciate ligament':ti,ab OR 'ligament':ti,ab)

CINAHL: (MH "Reproducibility of Results") OR (MH "Sensitivity and Specificity") OR (MH "Equipment Reliability") OR (MH "Reliability and Validity") OR (MH "Validation Studies") OR (MH "Instrument Validation") OR (MH "Precision") OR TI("validation" OR "validity" OR "accuracy" OR "accurately" OR "reliability" OR "reliable" OR "reproducible" OR "reproducibility" OR "repeatability" OR "precision") OR AB("validation" OR "validity" OR "accuracy" OR "accurately" OR "reliability" OR "reliable" OR "reproducible" OR "reproducibility" OR "repeatability" OR "precision") AND (MH "Arthrometry") OR (MH "Joint Instability/DI") OR (MH "Knee Injuries/DI") OR (MH "Orthopedic Equipment and Supplies") OR TI("arthrometer" OR "arthrometers" OR "device" OR "devices" OR "equipment" OR "laximetry") OR AB("arthrometer" OR "arthrometers" OR "device" OR "devices" OR "equipment" OR "laximetry") AND (MH "Knee Injuries") OR (MH "Anterior Cruciate Ligament") OR (MH "Ligament Injuries") OR (MH "Knee") OR (MH "Ligaments, Articular") OR (MH "Knee Joint") OR (MH "joint instability") OR TI("ACL" OR "anterior cruciate ligament" OR "ligament") OR AB("ACL" OR "anterior cruciate ligament" OR "ligament")

SPORTDiscuss: (DE "ARTICULAR ligaments" OR DE" CRUCIATE ligaments OR DE "PATELLAR tendon" OR DE "POSTERIOR cruciate ligament" OR DE "MEDIAL collateral ligament (Knee)" OR DE "ANTERIOR cruciate ligament" OR DE "ANTERIOR cruciate ligament injuries" OR DE "POSTERIOR cruciate ligament surgery" OR DE "POSTERIOR cruciate ligament injuries" OR DE "PATELLAR ligament surgery" OR DE "PATELLAR ligament injuries" OR DE "ANTERIOR cruciate ligament surgery" OR DE "KNEE dislocation" OR DE "KNEE injuries" OR "anterior cruciate ligament" OR ACL) AND (DE "ORTHOPEDIC apparatus" OR DE "EQUIPMENT & supplies" OR arthrometer OR arthrometry OR instrument OR laximeter OR laximetry) AND (DE "Diagnosis" OR DE "EXAMINATIONS – validity" OR validity OR validation OR accuracy OR accurate* OR reliability OR reliable OR precision OR reproducible OR reproducibility OR repeatab*)

Appendix B: Bias Assessments

Table B1 Bias Assessments of the Studies Discussing Reliability, Based on COSMIN Box 6[a]

Study	Design requirements	Statistical methods	Other	Final decision
Okazaki et al[20]	+	+	++	+
Tardy et al[21]	+	+	–	–
Hemmerich et al[22]	++	+	–	–
Carpenter et al[23]	++	—	–	—
Haughom et al[24]	+	++	–	–
Hoshino et al[25]	+	—	++	—
Kubo et al[26]	+	—	–	—
Labbe et al[46]	+	+	++	+
Kuroda et al[27]	+	+	–	–
Labbe et al[45]	+	—	++	—
Shultz et al[28]	++	++	–	–
Branch et al[47]	+	+	++	+
Beckley et al[29]	++	++	–	–
Tsai et al[30]	++	++	–	–
Lopomo et al[31]	+	++	++	+
Berruto et al[32]	+	—	–	—
Napier et al[48]	+	–	++	–
Vaidya et al[33]	+	++	–	–
Krause et al[44]	++	++	–	–
Alam et al[34]	++	++	–	–
McQuade et al[35]	++	—	–	—
Wroble et al[49]	+	—	–	—
Mills and Hull[36]	++	—	–	—
Chung et al[37]	++	++	++	++
Lorbach et al[38]	++	+	–	–
Neumann et al[39]	++	—	–	—
Lorenz et al[40]	++	++	–	–
Almquist et al[41]	++	++	–	–
Nascimento et al[42]	+	++	–	–
Kang et al[43]	++	+	–	–

[a]For the COSMIN bias assessments, all questions of the checklist were answered with either very good (++), adequate (+), doubtful (–), or inadequate (– –). The article's worst score defined the final score for each subsection and the overall final scores.

Table B2 Bias Assessments of the Studies Discussing the Validity, Based on COSMIN Box 8[a]

Study	Statistical methods	Other	Final decision
Okazaki et al[20]	++	++	++
Tashiro et al[50]	++	–	–
Okazaki et al[51]	++	++	++
Espregueira-Mendes et al[52]	++	++	++
Hoshino et al[25]	++	–	–
Kubo et al[26]	++	–	–
Labbe et al[46]	++	++	—
Labbe et al[45]	—	++	—
Moewis et al[53]	++	–	–
Mayr et al[54]	++	—	—
Bellitti et al[55]	—	–	—
Nakamura et al[59]	++	-	-
Helfer et al[56]	++	++	++
Napier et al[48]	++	++	++
Alam et al[34]	++	–	–
Mills and Hull[36]	—	–	—
Almquist et al[57]	++	–	–
Mayr et al[58]	++	–	–
Kang et al[43]	—	–	—

[a]For the COSMIN bias assessments, all questions of the checklist were answered with either very good (++), adequate (+), doubtful (–), or inadequate (– –). The article's worst score defined the final score for each subsection and the overall final scores.

Table B3 Bias Assessments of the Studies Discussing Diagnostic Accuracy, Based on QUADAS2[a]

Study	Patient selection	Index test	Reference standard	Flow and timing	Final decision
Tashiro et al[50]	+/–	+	+	+	+/–
Okazaki et al[51]	+/–	+/–	+	+	+/–
Espregueira-Mendes et al[52]	+	+	+	+	+
Araki et al[60]	–	–	–	+	–
Labbe et al[45]	+/–	+/–	+	+	+/–
Berruto et al[32]	+/–	–	+	+	–
Napier et al[48]	+	–	+	+	–
Vaidya et al[33]	+/–	–	+	+/–	–
Mouton et al[61]	+/–	–	+	+/–	–

[a]In QUADAS 2, bias assessment studies were evaluated to either have a low risk (+), a high risk (–), or the risk level was unclear (+/–). The lowest score of the domains defined the final score.