Injury Risk Estimation Expertise

Assessing the ACL Injury Risk Estimation Quiz

Erich J. Petushek,*†‡ PhD, CSCS, Edward T. Cokely,‡§ PhD,
Paul Ward,† PhD, CPsychol, CSci, AFBPsS, John J. Durocher,|| PhD, ACSM EP-C, CSCS,
Sean J. Wallace,¶ BS, and Gregory D. Myer,#**††‡‡ PhD, FACSM, CSCS*D
Investigation performed at Michigan Technological University, Houghton, Michigan, USA, and Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

Background: Available methods for screening anterior cruciate ligament (ACL) injury risk are effective but limited in application as they generally rely on expensive and time-consuming biomechanical movement analysis. A potential efficient alternative to biomechanical screening is skilled movement analysis via visual inspection (ie, having experts estimate injury risk factors based on observations of athletes' movements).

Purpose: To develop a brief, valid psychometric assessment of ACL injury risk factor estimation skill: the ACL Injury Risk Estimation Quiz (ACL-IQ).

Study Design: Cohort study (diagnosis); Level of evidence, 3.

Methods: A total of 660 individuals participated in various stages of the study, including athletes, physicians, physical therapists, athletic trainers, exercise science researchers/students, and members of the general public in the United States. The ACL-IQ was fully computerized and made available online (www.ACL-IQ.org). Item sampling/reduction, reliability analysis, cross-validation, and convergent/discriminant validity analysis were conducted to optimize the efficiency and validity of the assessment.

Results: Psychometric optimization techniques identified a short (mean time, 2 min 24 s), robust, 5-item assessment with high reliability (test-retest: r = 0.90) and consistent discriminability (average difference of exercise science professionals vs general population: Cohen d = 1.98). Exercise science professionals and general population individuals scored 74% and 53% correct, respectively. Convergent and discriminant validity was demonstrated. Scores on the ACL-IQ were most associated with ACL knowledge and various cue utilities and were least associated with domain-general spatial/decision-making ability, personality, or other demographic variables. Overall, 23% of the total sample (40% exercise science professionals; 6% general population) performed better than or equal to the ACL nomogram.

Conclusion: This study presents the results of a systematic approach to assess individual differences in ACL injury risk factor estimation skill; the assessment approach is efficient (ie, it can be completed in <3 min) and psychometrically robust. The results provide evidence that some individuals have the ability to visually estimate ACL injury risk factors more accurately than other instrument-based ACL risk estimation methods (ie, ACL nomogram). The ACL-IQ provides the foundation for assessing the efficacy of observational ACL injury risk factor assessment (ie, does simple *skilled* visual inspection reduce ACL injuries?). It also provides a representative task environment that can be used to increase our understanding of the perceptual-cognitive mechanisms underlying observational movement analysis and to improve injury risk assessment performance.

Keywords: ACL injury risk; injury prediction; movement analysis; expertise; psychometric; validation; reliability; test development

Female athletes are approximately 3 times more likely to rupture their anterior cruciate ligament (ACL) compared with their male counterparts.[39] Younger athletes (aged 15-25 years) participating in landing and cutting sports such as basketball and soccer are at greatest risk for noncontact ACL injury.[13] This elevated risk coupled with a nearly 2-fold increase in female sports participation over the past 30 years[20,34] has led to a rapid increase in the prevalence of ACL injuries in women (≈1 injury per 20 individuals during a sports season).[39] The cost of ACL surgery has been shown to be approximately US$5000, which does not include postoperative rehabilitation or lost time from work or sport.[33] In the United States alone, the annual cost of ACL injury likely exceeds US$3 billion.[21] Additional consequences of ACL injury include time out of sport or school, scholarship loss, risk for reinjury, and increased risk for osteoarthritis.[3,23,47] Interestingly, most ACL injuries in female athletes occur in a noncontact situation[1,22] and are likely preventable,[15] especially in young athletes.[32]

The American Journal of Sports Medicine, Vol. XX, No. X
DOI: 10.1177/0363546515580791

Neuromuscular training reduces the relative risk for noncontact ACL injury by 73.4%.[41] Unfortunately, 108 individuals must participate in training to prevent a single injury.[41] The time commitment involved in training this number of individuals is nontrivial. Moreover, athletes at high risk for ACL injury are more responsive to adaptations associated with prevention techniques including physical training.[28] Administering nontargeted prevention programs to low-risk athletes may reduce the efficiency and focus of time and resources. If training instead targeted high-risk athletes, the number of individuals needed to train to prevent an injury could be significantly reduced.

Methods for screening ACL injury risk factors have been identified and developed for high school–age (15-19 years) female athletes using prospective 3-dimensional (3D) biomechanical analysis procedures[14] as well as a more cost-effective nomogram approach.[30,31] The simplistic assessment approach used in the ACL nomogram represents considerable progress in the development of cost-effective, efficient screening tools. Central to the current investigation, the focus on assessment simplicity could, theoretically, be leveraged to produce even simpler, less expensive, yet valid assessment tools based on observational movement diagnostics.

Visual inspection or observational movement diagnosis is one alternative screening method that would reduce overall injury prevention time and cost.[i] Unfortunately, the accuracy and consistency of observational assessment of ACL injury risk are poorly understood. Previous research has demonstrated that physiotherapists[ii] are able to observationally assess specific variables purported to be associated with ACL injury risk.[7,35,40,44,45] However, these variables have not been supported by evidence from longitudinal prospective studies of ACL injury risk. In addition, the skill of observational injury risk estimation has not been assessed in other populations who may benefit from or commonly use this skill (eg, coaches, athletes, parents, athletic trainers, strength and conditioning coaches, medical doctors). Because individual differences in observational ACL injury risk estimation likely exist, a test quantifying ACL injury risk factor assessment skill could inform individuals about their level of proficiency and serve as a foundation for justifying the use of simple observation as a valid screening method. Therefore, the purpose of the current study was to develop a quick, robust, reliable, and valid test of ACL injury risk estimation skill: the ACL Injury Risk Estimation Quiz (ACL-IQ).

METHODS

To develop and validate the ACL-IQ, we performed item sampling, item reduction, reliability analysis, out-of-sample cross-validation, and convergent/discriminant validity analysis as outlined below.

Participants

Data were collected on a total of 660 individuals throughout the test development and validation process. Various exercise science professionals (eg, physical therapists/physiotherapists, athletic trainers, strength and conditioning coaches, exercise science students, sports medicine researchers, and physicians) participated in the "exercise science" group (n = 269). Additionally, individuals from the general population, such as sport coaches, athletes, parents, and others who are not exercise scientists, participated in the "general population" group (n = 391). The majority of the general population members were recruited from a paid webpanel service provided by Amazon's Mechanical Turk. This internet sample tends to be demographically diverse and more representative of the general US population than traditional university participants (see Paolacci et al[36]) but is slightly younger and more educated than the general population. This internet sample was limited to individuals located in the United States and an approval rating greater than 85% (ie, individual work must be approved by the requester before payment). Specific occupational/subgroup information about the participants is described in Table 1.

A summary of the various samples used in the test development and validation is depicted in Figure 1.

Procedures

The Cincinnati Children's Hospital Medical Center and Michigan Technological University Institutional Review Boards approved the study procedures. The study was fully computerized and hosted online. The study was not compatible with mobile devices and, therefore, was limited to laptop or desktop computers.

Item Sampling. Female athletes participating in landing and cutting sports are at the greatest risk for noncontact ACL injury compared with their male counterparts.[1,2,6,16]

*Address correspondence to Erich J. Petushek, PhD, CSCS, School of Human and Health Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK (email: erich.petushek@gmail.com).
†School of Human and Health Sciences, University of Huddersfield, Huddersfield, UK.
‡Department of Cognitive and Learning Sciences, Michigan Technological University, Houghton, Michigan, USA.
§Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany.
||Department of Biological Sciences, Michigan Technological University, Houghton, Michigan, USA.
¶Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois, USA.
#Division of Sports Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA.
**Departments of Pediatrics and Orthopaedic Surgery, University of Cincinnati, Cincinnati, Ohio, USA.
††The Sports Health and Performance Institute, OSU Sports Medicine, Ohio State University Medical Center, Columbus, Ohio, USA.
‡‡The Micheli Center for Sports Injury Prevention, Waltham, Massachusetts, USA.
One or more of the authors has declared the following potential conflict of interest or source of funding: Funding support was received from the National Science Foundation Graduate Research Fellowship (1108024), Research Council of Norway (230163), National Science Foundation (SES–1253263), and Ministerio de Economía y Competitividad (titled "Helping Doctors and Their Patients Make Decisions About Health," PSI2011-22954). This work was also supported in part by National Institutes of Health Grants R01-AR055563 and R01-AR056259.

TABLE 1
Participant Subgroups and Occupations (N = 660)

	n	% (of Group Total)
Exercise science group	269	
Athletic trainer	52	19
Physiotherapist/physical therapist	59	22
Physician[a]	39	14
Exercise science student	48	18
Exercise science academic	30	11
Strength and conditioning coach	41	15
General population group[b]	391	
Other	320	82
Parent of athlete	26	7
Young (≤25 y) female athlete	11	3
Sport coach	34	9

[a]82% of physicians specialized in orthopaedics/sports medicine and 18% in family medicine.

[b]Specific subgroup classification for 171 individuals from the general population group was not recorded, and therefore they are included in the "other" subgroup.

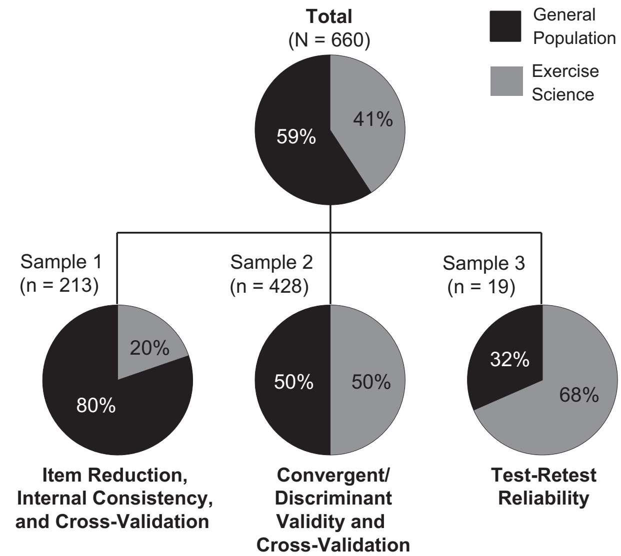


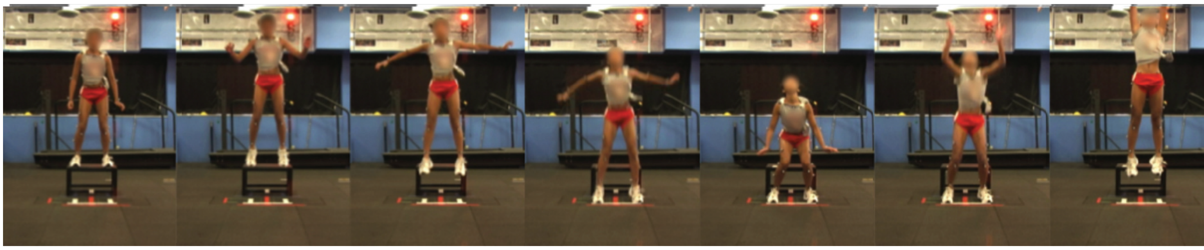
Figure 1. Summary of samples used to develop and validate the ACL Injury Risk Estimation Quiz (ACL-IQ). The total sample is depicted in the pie chart on top, followed by the specific subsamples and their specific use within the study.

Accordingly, initial items/stimuli consisted of a sample of 20 video clips of female athletes performing a drop vertical jump (Figure 2; see also the Video Supplement). The athletes featured in the videos participated in landing and cutting sports and also served as the participants for the development and validation of the clinical ACL nomogram[29] (mean ± SD: age, 15.9 ± 1.3 years; height, 163.6 ± 9.9 cm; body mass, 57 ± 12.1 kg). The athletes featured in the video stimuli were also demographically similar to individuals investigated in an initial prospective injury risk factor study (age, 16.0 ± 1.35 years; height, 165.9 ± 6.4; body mass, 60.3 ± 8.2 kg).[14] The 20 candidate video stimuli items also depicted athletes who have a wide and representative range of injury risk values (ie, knee abduction moments from very low to very high).

Risk Estimation. Participants viewed brief videotaped clips of athletes performing a drop vertical jump and were asked to estimate the risk for future ACL injury on a 10-point scale (Figure 2 and Video Supplement). The videotaped clips were approximately 3 seconds in length, and participants watched them only once (in real time) before being asked to assess injury risk. ACL injury risk was estimated using concurrent 3D biomechanical analysis of peak knee abduction moment,[30] which was linearly transformed on a 1 to 10 scale to quantify judgment accuracy. Individual items were scored by calculating the difference in score between the subjective rating and biomechanical criterion (ie, absolute value of the subjective rating score minus the biomechanical criterion score), thus representing judgment error. For example, if an individual rated the video clip as 4 "risk" and the criterion "risk" was 7, the error score was 3 (absolute value of (4 – 7) = 3).

Item Reduction. An initial sample of 213 individuals was used to select optimal items (171 in the general population and 42 in the exercise science groups). Based on constraints of the scale type, such as the 1 to 10 rating system, the current analysis used a hybrid approach to item analysis and reduction. Specifically, we used an item response theory–inspired, modified classical test theory item analysis (for a related approach using decision trees, see Cokely et al[4]). Our analyses examined individual item performance across 3 dimensions: difficulty, discriminability, and guessing. Difficulty was examined by calculating the average judgment error for each item across all individuals (eg, larger values indicated greater error and thus greater difficulty). Discriminability for each item was calculated as the standardized mean difference in judgment error between the exercise science professionals and general population individuals for each item. The guessing parameter is directly related to the item's known location on the scale (ie, the criterion injury risk category). The likelihood of obtaining a "correct" answer is greater for items near the middle of the scale (ie, 5). Thus, selecting items toward the ends of the scale should reduce the potential benefit of anchoring on any single rating. Guessing performance was analyzed using Monte Carlo simulations of 10,000 "pseudorandom" test performances. For each item in the test, an integer (1-10) was drawn from a standard normal distribution. Test performance was computed and averaged across the 10,000 iterations to estimate "chance" or guessing performance for each individual item and for overall test performance. Theoretically, the goal was to select items to represent a wide range of difficulty with a maximum degree of discriminability, while equally sampling from the full range of the scale to reduce artificial test score inflation resulting from any anchoring bias or guessing effect.

Six candidate tests ranging from 7 to 5 items (ie, video clips) were constructed and evaluated based on their various psychometric properties such as range, difficulty, discriminability, and guessing. All candidate tests were found to be sufficiently difficult—no individual attained perfect performance (ie, average peak performance across



	Very Low									Very High
Risk for ACL Injury	○ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10

Figure 2. Sample decision task. This sequence of snapshots from a video clip was shown to the participants in the decision task.

all candidate tests was 88%). The 7- and 6-item tests continued to display some moderate potential for anchoring bias such that an anchoring strategy (eg, selecting a risk rating of "5" for all items/clips) would result in better than chance performance. Thus, all 7- and 6-item tests were excluded from further consideration. Both 5-item tests reduced the test bias associated with the anchoring strategy (anchoring and chance were within 2 raw points or 6%). Additionally, both 5-item tests showed considerable similarity across guessing, anchoring, and central tendency variables (ie, guessing, 52%; anchoring, 58%; mean, 55%; median, 57%). Both 5-item tests also displayed similar difficulty values; however, the 5A-item test exhibited a 13% improvement in discriminability values compared with the 5B-item test (ie, a difference of 0.24 standard deviations). The 5A-item test also exhibited a wider range of scores despite maintaining similar discriminability values. The two 5-item tests differed on only 1 of the 5 items. Specifically, the 5A-item test included a more difficult item that replaced a moderately difficult item in the 5B-item test. Overall, results based on the hybrid item analysis indicated that the 5A-item test offered a highly desirable psychometric profile for all essential test-performance variables. Thus, the 5A-item test was chosen as the final ACL-IQ.

Reliability Analysis. Internal consistency was assessed on the 213 individuals previously used for item reduction (general population, n = 171; exercise science, n = 42) by quantifying judgment consistency error on repeated assessment of 5 different video clips (ie, the absolute value of the first subjective rating minus the second subjective rating of the same clip). In addition, test-retest reliability analysis was conducted by administering the final ACL-IQ to a separate sample of 19 individuals (sample 3) on 2 occasions separated by approximately 9 days. Test-retest correlation coefficient and typical error were quantified.[18,19]

Cross-validation. The initial sample used for test development was composed of a limited number of exercise science/sports medicine personnel (ie, 42), resulting in some risk for sample and thus test bias. Therefore, after the initial sample of 213 individuals used to select optimal items (sample 1: 171 general population and 42 exercise science), an additional 428 individuals (sample 2: 214 general population and 214 exercise science) completed the ACL-IQ. Between-group discriminability results (Cohen *d*), peak performance, and range were compared between samples to ensure that the ACL-IQ demonstrated consistent discriminability and difficulty in a larger representative sample.

Convergent/Discriminant Validity. To the extent that the ACL-IQ measures an acquired skill at predicting ACL injury risk, rather than a more general ability, test performance should be related to domain-specific factors such as ACL knowledge (ie, knowledge about location, function, and risk factors for ACL injury) and specific strategies or cues used in risk assessment such as medial knee motion, landing stiffness, and limb asymmetry. Test performance should not be related to domain-general measures of mental (ie, cognitive) capacity such as general intelligence, decision-making ability, or spatial reasoning ability. After completing the ACL-IQ, participants answered 11 ACL knowledge questions related to location, function, and risk factors for injury (see the Appendix). Additionally, the perceived utility of specific information cues that might aid in a participant's risk assessment, henceforth, cue utility, was elicited through a brief survey in which participants rated (on a 1-10 scale) the importance of specific visual cues (ie, knee motion, hip motion, trunk motion, landing stiffness, height, weight (see the Appendix, available online at http://ajsm.sagepub.com/supplemental). Subjective rating of cue utility has been used in other expertise studies such as the feature discrimination task.[24-26,46]

To estimate the potential relationship of general cognitive abilities to overall test performance and, specifically, to assess discriminant validity, participants completed the Berlin Numeracy Test (BNT). The BNT has been extensively validated for assessment of statistical numeracy and risk literacy, which is the ability to accurately interpret and make good decisions based on information about risk.[4] Theoretically, it is also possible that domain-general spatial abilities help determine observational movement analysis performance (eg, ACL-IQ performance). Hence, the 24-item Mental Rotation Test (MRT-A) was administered to examine this possibility.[37,42] Finally, to add further discriminant validity evidence and to examine potential test bias related to ACL-IQ scores, personality was assessed using the 10-item Big Five.[12]

Clinical Standard Comparison. The ACL-IQ scores can be transformed into practical meaning by simply subtracting the score from 38 (maximum points) and dividing by 5 (number of items per video clip). This value represents the average deviation from the criterion on any given video clip or test item. For example, if an individual scored a 28, his or her average absolute error would be 2.00 ([38 – 28]/5), meaning if a video clip was presented with a criterion risk value of 5 (on the 1-10 point scale) (Figure 2), this individual would, on average, be within ±2 of the criterion or between 3 and 7. A mean error of 2 may seem unacceptable to some, but if the purpose of identifying the ACL injury risk level of an athlete (ie, screening) is to decide on an appropriate intervention (eg, feedback, training, etc), the athlete may only need to be classified into a "high" or "low" risk group. Given the small number of items (ie, 5), an alternative to standard signal detection analysis can be used to compare ACL-IQ scores with high/low risk classification accuracy. If we classify the risk level of an athlete at greater than 5 (on the 10-point scale) as "high" risk, we can determine the number of judges who correctly classified all 5 athletes (video clips/test items) into either "high" (ie, above 5) or "low" (ie, below 5) risk categories.[iii] The ACL nomogram can also produce an ACL-IQ score, which can be used as a clinical benchmark for comparison with an individual's observational judgments.[iv]

RESULTS

Reliability Analysis

Initial internal consistency results revealed that the majority of individuals showed a high level of consistency in their judgments (ie, the absolute consistency error was 1.09), although exercise science professionals showed slightly higher levels of consistency (mean absolute consistency error ± SD: exercise science professionals, 0.83 ± 0.45; general population, 1.15 ± 0.78; Cohen d = 0.45). Specifically, for a repeated trial, the average expected deviation on the second rating was 0.83 (on a 10-point rating scale) for the exercise science group. An absolute consistency error value ≤1.00 will likely not influence injury risk estimation if fewer categories were used (ie, from 10 to 3) which may best represent decision or intervention points. Moreover, 76% of the exercise science professionals displayed absolute consistency error values at or below 1.00, compared with 54% of the general population group.

Table 2 describes the test-retest characteristics of the ACL-IQ. Despite the high test-retest correlation coefficient (r = 0.90), a small mean difference (1.53 points) was displayed between test sessions. The mean difference was small, within the 90% typical error range (0-2.11) and similar to internal consistency estimates in a previous sample of 213 (1.09). The typical error represents the amount the score may vary on a repeated test performance. For example, if someone scored 30 out of 38 (ie, 8 error points) on the first ACL-IQ, there is a 90% probability that his or her score on a repeated performance would be between 28 and 32 (typical error × 1.65 = 2.11). Thus, based on this typical error profile, it is highly unlikely that this statistically significant difference of 1.53 is clinically meaningful.

TABLE 2
ACL-IQ Test-Retest Reliability Characteristics (n = 19)[a]

	Test	Retest
Descriptive statistics		
Mean score (%)	28.47 (75)	26.95 (71)
SD (%)	3.82 (10)	4.10 (11)
Range (%)	20-34 (53-89)	18-33 (47-87)
Time between tests, d, mean ± SD	9.42 ± 2.78	
Reliability metrics		
Test-retest correlation (95% CI[b])	0.90 (0.74-0.96)	
Typical error (%)	1.28 (3)	
Mean difference (%)	–1.53 (–4)[c]	
Cohen d	0.39	

[a]% indicates percentage correct. ACL-IQ, ACL Injury Risk Estimation Quiz.
[b]Bootstrap using 1000 samples.
[c]Significant mean difference, $t(18) = -3.68$, $P = .002$.

Cross-validation

Range and average measures of ACL-IQ scores from sample 1 (171 [80%] general population and 42 [20%] exercise science), and sample 2 (214 [50%] general population and 214 [50%] exercise science) are presented in Table 3. No statistically significant difference in effect size (Cohen d) was noted between the 2 samples (z = 1.17; P = .24). The average score including all individuals in sample 1 was 63%, which was statistically different from all individuals in sample 2, likely due to the larger number of exercise science professionals in sample 2 (50% vs 20%). Similar to sample 1, no individual in sample 2 scored 100% correct. Additionally, subgroup (exercise science and general population) means between samples 1 and 2 were nearly identical, corroborating the discriminability evidence for the ACL-IQ.

Convergent/Discriminant Validity

Initial independent correlations between various factors and ACL-IQ are displayed in Table 4. The factors displaying large associations with ACL-IQ performance (ie, $r \geq 0.40$) were ACL knowledge, education level, and cue utility variables knee/thigh motion and jump height. Significant independent task-relevant cues included inward/outward knee/thigh motion and lateral trunk motion. Significant task-irrelevant cues included height and weight of the individual as well as jump height and jump alignment. Domain-general perceptual-cognitive abilities, although statistically significant, remained marginally associated with risk estimation performance (ie, r values < 0.25). Additionally, personality characteristics displayed minimal or no association with ACL-IQ.

Clinical Standard Comparison

Overall, 20% of the total sample classified all 5 video clips into correct "high" and "low" risk categories. Group-wise, 35% of exercise science professionals and 4% of the general

TABLE 3
Sample 1 and 2 Cross-validation Comparison[a]

Scale Attributes	Sample 1 (n = 213)	Sample 2 (n = 428)
Time, min:s, mean ± SD	2:24 ± 0:47	
Score range	0-38 ± 0-100	
Achieved range (%)	12-34 (32-89)	10-36 (26-95)
Overall mean (%)	21.69 (57)	24.00 (63)[b]
Overall median (%)	21 (55)	24 (63)
Overall SD (%)	5.11 (13)	5.86 (15)
Discriminability		
Exercise science		
n (% of sample)	42 (20)	214 (50)
Mean (%)	28.31 (74)[c]	27.97 (74)[c]
SD (%)	3.80 (10)	3.97 (10)
General population		
n (% of sample)	171 (80)	214 (50)
Mean (%)	20.07 (53)	20.04 (53)
SD (%)	3.96 (10)	4.63 (12)
Cohen d	2.11	1.84
Weighted SD (%)	3.92 (10)	4.30 (11)
95% CI	1.70-2.48	1.60-2.05

[a]% indicates percentage correct.
[b]Significantly different from sample 1, $P < .01$.
[c]Significantly different from general population group, $P < .01$.

population group classified all 5 video clips into correct "high" and "low" risk categories. The average ACL-IQ for these individuals with 100% 2-category classification accuracy was 31.18 (6.81 error points, mean error of 1.34, or 82% correct). When transformed into 1 to 10 categories, the ACL nomogram demonstrated 8 error points (ACL-IQ score of 30, or 79% correct). Overall, 23% of the total sample performed better than or equal to the ACL nomogram. Group-wise, 40% of exercise science professionals and 6% of the general population group performed better than or equal to the ACL nomogram.

DISCUSSION

This investigation developed and assessed the psychometric properties of a test designed to evaluate observational ACL injury risk factor estimation ability (ACL-IQ). The results revealed that the ACL-IQ is an efficient, robust, and reliable research tool. Additionally, convergent/discriminant validity evidence was established demonstrating that the ACL-IQ assessment works well because it conforms with current theories of expertise where domain-specific factors and, importantly, judgment processes (eg, cue usage) contribute highly to describe superior performance.[9-11,27] A more comprehensive analysis of the underlying performance mechanisms using structural and path (eg, mediation) modeling is warranted and will likely reveal that the demographic and domain-general factors associated with ACL-IQ will be mediated by ACL knowledge and cue utilities (E.J. Petushek, E.T. Cokely, P. Ward, G.D. Myer, unpublished data, 2014).[38] For example, the large association between education level and ACL-IQ is likely a function of greater ACL knowledge and cue utilization. Understanding the perceptual-cognitive mechanisms of performance will be important for developing training. Furthermore, cognitive process tracing methods such as verbal protocol analysis or eye-tracking can be used to reverse-engineer superior performance to optimize training and decision support tools.[4,5,8,9,17,43]

TABLE 4
Independent Correlation With ACL-IQ (n = 428)[a]

	ACL-IQ	95% CI[b]
Domain-specific factors		
ACL knowledge test (11-items)	0.59[c]	0.54 to 0.65
ACL papers and books read per mo	0.38[c]	0.31 to 0.44
ACL risk assessment experiences, y	0.19[c]	0.11 to 0.28
Estimated cue validity (cue utility)		
Arm motion	–0.04	–0.13 to 0.05
Landing symmetry	0.08	–0.03 to 0.18
Inward/outward knee motion	0.40[c]	0.32 to 0.47
Inward/outward thigh motion	0.34[c]	0.26 to 0.42
Knee and thigh composite average[d]	0.40[c]	0.33 to 0.47
Lateral trunk motion	0.19[c]	0.1 to 0.29
Landing stiffness	0.01	–0.09 to 0.09
Foot alignment	–0.07	–0.16 to 0.02
Height of individual	–0.19[c]	–0.28 to –0.09
Weight of individual	–0.38[c]	–0.46 to –0.29
Jump height	–0.54[c]	–0.61 to –0.46
Jump alignment	–0.18[c]	–0.28 to –0.09
Domain-general factors		
Perceptual-cognitive ability		
Mental Rotation Test-A (24 items)[e]	0.24[c]	0.15 to 0.33
Berlin Numeracy Test (4 items)[e]	0.14[c]	0.05 to 0.24
Personality traits		
Extraversion	0.12[f]	0.03 to 0.23
Agreeableness	–0.11[f]	–0.21 to –0.01
Conscientiousness	0.17[c]	0.06 to 0.27
Emotional stability	0.06	–0.03 to 0.15
Openness to experience	–0.05	–0.14 to 0.05
Demographic variables		
Education level	0.40[c]	0.32 to 0.47
Age	–0.19[c]	–0.27 to –0.10
Sex	0.18[c]	0.09 to 0.27
Sport participation	0.30[c]	0.21 to 0.39
ACL injury diagnosis	0.13[c]	0.03 to 0.22

[a]ACL, anterior cruciate ligament; ACL-IQ, ACL Injury Risk Estimation Quiz.
[b]Bootstrap using 1000 samples.
[c]$P < .01$.
[d]Variable computed to replace both knee and thigh motion to decrease multicollinearity [knee and thigh motion: $r(427) = 0.71$].
[e]Mental Rotation Test (MRT) was missing 36 (8.4%) values and Berlin Numeracy Test (BNT) 8 (1.9%). Little's Missing Completely at Random (MCAR) test was not significant ($P = .53$); thus, expectation maximization (implemented in SPSS; SPSS Inc) was used to interpolate missing values and calculate the correlation coefficients, which were not statistically different from the coefficients with missing values [missing data MRT: $r(391) = 0.23$ (0.14-0.33), $P < .01$; and BNT: $r(419) = 0.14$ (0.05-0.24), $P < .01$]. Additionally, the means of the interpolated and missing datasets were not statistically different ($P < .01$).
[f]$P < .05$.

The collective clinical standards comparisons suggest that an ACL-IQ score of around 80% correct may be suitable for justifying the use of observation as a suitable screening method for ACL injury risk factor estimation. Moreover, 28% of the exercise science professionals and 3% of the general population had an ACL-IQ score at or above 80%. However, more data are needed to justify using observational screening, and, in particular, an appropriate signal detection analysis would reveal estimates of sensitivity, specificity, discriminability, and response bias that can be used to assess the efficacy of a screening approach (in addition to cost and time associated with the screening method and misses/false alarms).

Since there are limited data regarding the biomechanical risk factors for ACL injury, there is a great opportunity for future research to begin to understand whether skilled observation can be a suitable method for screening. Additional studies, prospective in design, are needed to assess the predictive validity for using the ACL-IQ to identify individuals who can predict ACL injury risk with suitable accuracy. A prospective injury risk study could be conducted by incorporating observational screening with appropriate training intervention and comparing injury rates with no screening or training everyone. Prospective studies are resource intensive and often require many years of data collection. To begin to understand whether observation can be used to assess ACL injury risk factors, a pseudoprospective study could be conducted using videotaped individuals (ie, drop vertical jump) who later went on to injure their ACL. Specifically, video clips of a representative sample of athletes could be shown to and rated by observers with various levels of ACL-IQ. Classification accuracy could then be established by comparing the observer ratings to actual outcomes (no injury/injury). This design would significantly reduce time and any ethical dilemmas associated with identifying injury risk level by unskilled individuals as well any confounding effects due to training. The goal would be to establish evidence that ACL-IQ scores are correlated with observers' classification accuracy (area under the receiver operating characteristic curve, sensitivity, specificity, etc) with actual injurious events.

CONCLUSION

The ACL-IQ is the first technology to assess individual differences in observational ACL injury risk factor estimating performance. This efficient tool demonstrated excellent psychometric properties. This work provides a foundation for future research investigating the degree to which simple observational screening can prevent ACL injuries. Further understanding of the underlying mechanisms of ACL-IQ can be used to develop training applications. The web-based platform at www.ACL-IQ.org enhances outreach and awareness and provides individualized feedback to professionals and the public. ACL-IQ.org is also a repository for continuous data collection and the future home of efficient training programs and decision support tools.

ACKNOWLEDGMENT

The authors thank the Oslo Sports Trauma Research Center Team (especially Tron Krosshaug) for data collection assistance. The authors also thank all of those who volunteered their time to participate in this research project and those who helped distribute the web-based survey (especially Kim Barber Foss and Cheri Baumann).

A Video Supplement for this article is available in the online version or at http://ajsm.sagepub.com/supplemental.

NOTES

i. With an injury cost of US$17,000, prevention training overall injury risk of 0.33%, estimated training time of 30 minutes per session for 32 sessions, individual training cost of $50, screening test sensitivity of 78% and specificity of $73%, and visual screening time of 1 minute per individual, the overall cost per individual would reduce by 28% and overall prevention time would reduce by 75% if visual screening plus training was implemented compared with training everyone (without screening).

ii. There are subtle differences between a physical therapist and physiotherapist with regard to name ownership, professional organizations, and accreditation. ''Physical therapist'' is the common title held in the US, whereas ''physiotherapist'' is the common title in UK/Australia and Canada and is likely the more recognized international term for this type of qualification. This study includes data from both European physiotherapists and US physical therapists.

iii. Three of the 5 ACL-IQ items/video clips demonstrated high risk (knee abduction moments $>$41 N·m) and 2 demonstrated low risk (knee abduction moments $<$17 N·m). Previous research has used a knee abduction moment of 25 and 22 N·m as a cut-point for high and low risk, respectively.

iv. The video clips used in this study also had concurrent ACL nomogram assessment for only the left leg. Right and left leg knee abduction moment (estimated risk criterion) demonstrated a correlation coefficient of $r(19) = 0.62$.

REFERENCES

1. Agel J, Arendt EA, Bershadsky B. Anterior cruciate ligament injury in national collegiate athletic association basketball and soccer: a 13-year review. *Am J Sports Med*. 2005;33(4):524-530.
2. Agel J, Olson DE, Dick R, Arendt EA, Marshall SW, Sikka RS. Descriptive epidemiology of collegiate women's basketball injuries: National Collegiate Athletic Association Injury Surveillance System, 1988-1989 through 2003-2004. *J Athl Train*. 2007;42(2):202.
3. Ardern CL, Webster KE, Taylor NF, Feller JA. Return to the preinjury level of competitive sport after anterior cruciate ligament reconstruction surgery: two-thirds of patients have not returned by 12 months after surgery. *Am J Sports Med*. 2011;39(3):538-543.
4. Cokely ET, Galesic M, Schulz E, Ghazal S, Garcia-Retamero R. Measuring risk literacy: the Berlin Numeracy Test. *Judgm Decis Mak*. 2012;7(1):25-47.
5. Cokely ET, Kelley CM. Cognitive abilities and superior decision making under risk: a protocol analysis and process model evaluation. *Judgm Decis Mak*. 2009;4(1):20-33.
6. Deitch JR, Starkey C, Walters SL, Moseley JB. Injury risk in professional basketball players: a comparison of Women's National Basketball Association and National Basketball Association athletes. *Am J Sports* Med. 2006;34(7):1077-1083.
7. Ekegren CL, Miller WC, Celebrini RG, Eng JJ, Macintyre DL. Reliability and validity of observational risk screening in evaluating dynamic knee valgus. *J Orthop Sports Phys Ther*. 2009;39(9):665-674.

8. Ericsson KA, Charness N, Feltovich PJ, Hoffman RR. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, UK: Cambridge University Press; 2006.
9. Ericsson KA, Krampe RT, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev*. 1993;100(3):363-406.
10. Ericsson KA, Lehmann AC. Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu Rev Psychol*. 1996;47:273-305.
11. Ericsson KA, Prietula MJ, Cokely ET. The making of an expert. *Harv Bus Rev*. 2007;85(7/8):114.
12. Gosling SD, Rentfrow PJ, Swann WB Jr. A very brief measure of the Big-Five personality domains. *J Res Pers*. 2003;37(6):504-528.
13. Griffin LY, Albohm MJ, Arendt EA, et al. Understanding and preventing noncontact anterior cruciate ligament injuries: a review of the Hunt Valley II meeting, January 2005. *Am J Sports Med*. 2006;34(9):1512-1532.
14. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med*. 2005;33(4):492-501.
15. Hewett TE, Myer GD, Ford KR, Paterno MV, Quatman CE. The 2012 ABJS Nicolas Andry Award: the sequence of prevention: a systematic approach to prevent anterior cruciate ligament injury. *Clin Orthop Relat Res*. 2012;470(10):2930-2940.
16. Hewett TE, Zazulak BT, Krosshaug T, Bahr R. Clinical basis: epidemiology, risk factors, mechanisms of injury, and prevention of ligament injuries of the knee. In: Bonnin M, Amendola NA, Bellemans J, MacDonald SJ, Menetrey J, eds. *The Knee Joint: Surgical Techniques and Strategies*. Paris, France: Springer-Verlag; 2012: 53-70.
17. Hoffman RR, Ward P, Feltovich PJ, DiBello L, Fiore SM, Andrews DH. *Accelerated Expertise: Training for High Proficiency in a Complex World*. New York, NY: Guilford Press; 2014.
18. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30(1):1-15.
19. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc*. 2009;41(1):3-13.
20. Irick E. NCAA sports sponsorship and participation rates report 1981-1982–2011-1012. http://www.ncaapublications.com/productdownloads/PR2013.pdf. Published 2012.
21. Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of national surveys of ambulatory surgery, 1996 and 2006. *J Bone Joint Surg Am*. 2011;93:994-1000.
22. Krosshaug T, Nakamae A, Boden BP, et al. Mechanisms of anterior cruciate ligament injury in basketball: video analysis of 39 cases. *Am J Sports Med*. 2007;35(3):359-367.
23. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. *Am J Sports Med*. 2007;35(10):1756-1769.
24. Loveday T, Wiggins MW, Harris JM, O'Hare D, Smith N. An objective approach to identifying diagnostic expertise among power system controllers. *Hum Factors*. 2013;55(1):90-107.
25. Loveday T, Wiggins MW, Searle BJ. Cue utilization and broad indicators of workplace expertise. *J Cogn Eng Decis Making*. 2013;8(1):98-113.
26. Loveday T, Wiggins MW, Searle BJ, Festa M, Schell D. The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Hum Factors*. 2012;55(1):125-137.
27. Mann DTY, Williams AM, Ward P, Janelle CM. Perceptual-cognitive expertise in sport: a meta analysis. *J Sport Exerc Psychol*. 2007;29:457-478.
28. Myer GD, Ford KR, Brent JL, Hewett TE. Differential neuromuscular training effects on ACL injury risk factors in ''high-risk'' versus ''low-risk'' athletes. *BMC Musculoskelet Disord*. 2007;8:39-46.
29. Myer GD, Ford KR, Hewett TE. New method to identify athletes at high risk of ACL injury using clinic-based measurements and freeware computer analysis. *Br J Sports Med*. 2011;45(4):238-244.
30. Myer GD, Ford KR, Khoury J, Hewett TE. Three-dimensional motion analysis validation of a clinic-based nomogram designed to identify high ACL injury risk in female athletes. *Phys Sportsmed*. 2011;39(1):19-28.
31. Myer GD, Ford KR, Khoury J, Succop P, Hewett TE. Development and validation of a clinic-based prediction tool to identify female athletes at high risk for anterior cruciate ligament injury. *Am J Sports Med*. 2010;38(10):2025-2033.
32. Myer GD, Sugimoto D, Thomas S, Hewett TE. The influence of age on the effectiveness of neuromuscular training to reduce anterior cruciate ligament injury in female athletes: a meta-analysis. *Am J Sports Med*. 2013;41(1):203-215.
33. Nagda SH, Altobelli GG, Bowdry KA, Brewster CE, Lombardo SJ. Cost analysis of outpatient anterior cruciate ligament reconstruction: autograft versus allograft. *Clin Orthop Relat Res*. 2010;468(5):1418-1422.
34. NFSHSA. 2011-12 high school athletics participation survey. http://www.nfhs.org/ParticipationStatics/PDF/2011-12%20Participation%20Survey.pdf. Published 2012.
35. Nilstad A, Andersen TE, Kristianslund E, et al. Physiotherapists can identify female football players with high knee valgus angles during vertical drop jumps using real-time observational screening. *J Orthop Sports Phys Ther*. 2014;44(5):358-365.
36. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. *Judgm Decis Mak*. 2010;5(5):411-419.
37. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain Cogn*. 1995;28(1):39-58.
38. Petushek EJ. *Development and Validation of the Anterior Cruciate Ligament Injury-Risk-Estimation Quiz (ACL-IQ)*. Houghton, MI: Applied Cognitive Science and Human Factors, Michigan Technological University; 2014.
39. Prodromos CC, Han Y, Rogowski J, Joyce B, Shi K. A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sport, and a knee injury-reduction regimen. *Arthroscopy*. 2007;23(12):1320-1325.
40. Stensrud S, Myklebust G, Kristianslund E, Bahr R, Krosshaug T. Correlation between two-dimensional video analysis and subjective assessment in evaluating knee control among elite female team handball players. *Br J Sports Med*. 2010;45(7):589-595.
41. Sugimoto D, Myer GD, McKeon JM, Hewett TE. Evaluation of the effectiveness of neuromuscular training to reduce anterior cruciate ligament injury in female athletes: a critical review of relative risk reduction and numbers-needed-to-treat analyses. *Br J Sports Med*. 2012;46(14):979-988.
42. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. *Percept Mot Skills*. 1978;47:599-604.
43. Ward P, Suss J, Basevitch I. Expertise and expert performance-based training (ExPerT) in complex domains. *Technol Instr Cognition Learn*. 2009;7(2):121-145.
44. Whatman C, Hing W, Hume P. Physiotherapist agreement when visually rating movement quality during lower extremity functional screening tests. *Phys Ther Sport*. 2012;13(2):87-96.
45. Whatman C, Hume P, Hing W. The reliability and validity of physiotherapist visual rating of dynamic pelvis and knee alignment in young athletes. *Phys Ther Sport*. 2013;14(3):168-174.
46. Wiggins M, Brouwers S, Davies J, Loveday T. Trait-based cue utilization and initial skill acquisition: implications for models of the progression to expertise. *Cognition*. 2014;5(541):1-8.
47. Wright RW, Dunn WR, Amendola A, et al. Risk of tearing the intact anterior cruciate ligament in the contralateral knee and rupturing the anterior cruciate ligament graft during the first 2 years after anterior cruciate ligament reconstruction: a prospective MOON cohort study. *Am J Sports Med*. 2007;35(7):1131-1134.