Unilateral tests of lower-limb function as prognostic indicators of future knee-related outcomes following anterior cruciate ligament injury: a systematic review and meta-analysis of 13 150 adolescents and adults

Thomas J West 1,2, Andrea M Bruder 1,2, Kay M Crossley 1,2, Adam G Culvenor 1,2

ABSTRACT

Objective To investigate the prognostic capacity of individual hop tests, hop test batteries and other unilateral functional performance tests following anterior cruciate ligament (ACL) injury.

Design Systematic review with meta-analysis.

Data sources Six databases searched up to June 2021.

Eligibility criteria Studies reporting associations between unilateral lower-limb function (eg, hop tests) following ACL injury and future (≥3 months) knee-related outcomes.

Results Of 42 included studies (13 150 participants), all assessed the single-forward hop test and 32 assessed a repeated-forward hop test (crossover hop, triple hop, 6m-timed hop), mostly within a year after ACL injury/reconstruction. Results of meta-analyses indicated that higher single-forward and repeated-forward hop limb symmetry were associated with higher odds of return-to-sport 1–3 years post-ACL reconstruction (OR 2.15; 95% CI 1.30 to 3.54; OR 2.11; 95% CI 1.23 to 3.40, respectively). Higher single-forward and repeated-forward hop limb symmetry was associated with better self-reported symptoms and function 1–37 years after ACL injury (OR 2.51; 95% CI 1.62 to 3.88; OR 4.28; 95% CI 1.65 to 11.08, respectively). Higher limb symmetry on a repeated-forward hop does not appear to be associated with higher odds of successful rehabilitation without ACL reconstruction (OR 1.51; 95% CI 0.94 to 2.44). Achieving ≥90% limb symmetry on the single-forward hop was associated with reduced odds of knee osteoarthritis 5–37 years after ACL injury (OR 0.46; 95% CI 0.23 to 0.94).

Conclusion Very low certainty evidence suggests single-forward and repeated-forward hop tests are prognostic indicators for important knee-related outcomes in individuals after ACL injury and may help stratify individuals at risk of poor outcomes to target rehabilitation interventions.

PROSPERO registration number CRD42018092197.

WHAT IS ALREADY KNOWN?

⇒ Unilateral measures of lower-limb function are recommended for return to sport clearance after anterior cruciate ligament (ACL) injury.

⇒ Hop test results are important indicators of functional recovery after ACL injury.

WHAT ARE THE NEW FINDINGS?

⇒ Very low certainty evidence indicates that a higher Limb Symmetry Index for hop tests (irrespective of specific hop test assessed) is a prognostic indicator for returning to competitive sport but may not infer a reduction of future injury risk.

⇒ Very low certainty evidence indicates that higher Limb Symmetry Index on single-forward, repeated-forward hop tests and the one-leg rise test are prognostic for better self-reported symptoms and function 1–37 years after ACL injury.

⇒ Very low certainty evidence indicates that a battery of hop tests is associated with reduced odds of knee reinjury though sensitivity analysis revealed conflicting findings with different test batteries.

⇒ Very low certainty evidence indicates that achieving ≥90% Limb Symmetry Index on a single-forward hop test between 6 months and 4 years after ACL injury were associated with reduced odds of future knee osteoarthritis.

INTRODUCTION

Anterior cruciate ligament (ACL) rupture is a devastating diagnosis for athletes given the lengthy rehabilitation, absence from sports participation and high risk of early-onset osteoarthritis (OA).1 The primary target of rehabilitation, irrespective of operative or non-operative management, is addressing lower-limb muscle strength, neuromuscular control and functional deficits that can otherwise persist for many years.2 Functional performance testing is recommended during ACL injury rehabilitation to assess lower-limb function and guide return to sport.3 4 Functional performance measures also provide valuable feedback to enhance patient motivation and exercise adherence,5 6 integral to achieving desired outcomes following ACL injury and reconstruction.7

Hop tests are common performance measures used during ACL injury and reconstruction rehabilitation8 and given that they require few resources and are quick to administer, are widely used clinically. Many hop tests have been described (eg, hop for distance (single-forward hop), triple hop, crossover hop), with all having excellent inter-rater reliability.9–11 Hop tests underpin return to sport
test batteries and meeting acceptable standards (eg, ≥90% Limb Symmetry Index (LSI)) is a goal of rehabilitation.3

Some findings suggest that lower-limb functional test scores assessed during and following ACL rehabilitation are associated with future knee-related outcomes.12–14 For example, poor hop test performance is associated with worse future quality of life and fewer one-leg rises is related to development of radiographic OA.15 16 However, other studies found no association between functional performance and future outcomes.17–19 Systematic review evidence is also conflicting. Passing a return to sport test following ACL injury and identifying those at risk for worse outcomes to implement preventive strategies. This systematic review aimed to investigate the prognostic capacity of individual outcomes to implement preventive strategies. This systematic review aimed to investigate the prognostic capacity of individual hop tests, hop test batteries and other unilateral functional performance tests on knee-related outcomes following ACL injury.

**METHODS**

**Search strategy and selection criteria**

This systematic review is reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement24 and registered prospectively on PROSPERO (CRD42018092197). We performed an electronic search of six databases (MEDLINE, EMBASE, CINAHL, Scopus, Web of Science, SPORTDiscus) without language restriction from inception to June 2021. The searches combined MeSH terms and keywords related to ACL and hop tests (online supplemental appendix 1). Studies reporting the association (or lack thereof) between unilateral lower-limb functional performance following ACL injury and future outcomes were eligible for inclusion. Specifically, included studies had to: (1) report on individuals after ACL injury treated operatively or non-operatively (no limits to age or type of surgery); (2) include a unilateral functional test (quantitative measure); (3) able to be assessed in a clinical setting (ie, no laboratory or video-recording equipment needed) and (4) report the outcome ≥3 months following functional testing. We also included studies that assessed a hop test battery (ie, multiple hops combined into one overarching assessment). As it is common for hop test batteries to incorporate a measure of muscle strength,25 26 we also included studies that evaluated a hop test battery with muscle strength assessment as part of that battery. Cohort studies (prospective and retrospective), case-control studies and randomised controlled trials (RCTs) were eligible for inclusion. Any knee-related outcomes were considered (eg, patient-reported, return to sport, knee reinjury, OA).

Two reviewers (TJW and AGC) independently assessed all titles and abstracts for eligibility and screened reference lists of all relevant articles identified. When eligibility could not be confirmed from title and abstract, full texts were reviewed. Disagreements between reviewers were discussed until consensus, as described previously.25 During screening of titles, abstracts and full texts, articles published in languages other than English were excluded.

**Data extraction**

Two reviewers (TJW and AGC) independently extracted data using a predefined excel spreadsheet. Data extracted included participant characteristics (eg, age, sex, body mass index), functional test(s) assessed, outcome measures and results of functional tests and knee-related outcome. We contacted study authors for additional information as required.

**Risk of bias and level of evidence**

Two reviewers (TJW and AMB) independently assessed risk of bias using a modified version of the Newcastle-Ottawa scale.28 As we expected most eligible studies to be observational, we a priori developed a modified Newcastle-Ottawa scale for this review, consisting of ten items across four constructs (online supplemental appendix 2). The four constructs were: (1) participant selection; (2) definition of exposure; (3) comparability (exposed and nonexposed groups and (4) outcome assessment. Consistent with Cochrane Handbook for Systematic Reviews recommendations,29 each item was ranked as high or low risk of bias. Grade of Recommendations Assessments, Development and Evaluation (GRADE) assessed overall certainty of evidence.30 As part of the GRADE tool, where meta-analysis of any outcome included≥10 studies, publication bias secondary to small study effects was assessed using funnel plots and the Eggers test.

**Data synthesis and analysis**

Unilateral functional test data was calculated as an LSI defined as ACL injured limb performance ± contralateral limb performance × 100. We analysed LSI primarily as a continuous variable (mean and SD) but also explored the influence of dichotomised LSI results (ie, ≥90% limb symmetry) as this is a common threshold to define success.25 Knee-related outcome data were dichotomised to define an unsuccessful vs successful outcome using established criteria (table 1). These criteria were those most commonly used in the included studies (online supplemental appendix 3). Where results were not reported using these criteria, authors were contacted for data set out using these thresholds to facilitate meta-analysis.

Based on findings in a recent systematic review,31 we expected the single-forward hop to be most commonly assessed with sufficient data available for a meta-analysis. Given the similarity of repeated-forward hops (ie, triple hop, crossover hop and/or 6m-timed hop) we grouped these hops into one ‘repeated-forward hop’ category. Crossover hop results were preferred as it comprises forward and lateral movements and is comparable to others for reliability.32 If data from a crossover hop was not available, we used (in hierarchical order): (1) triple hop, (2) 6m-timed hop. To assess if one repeated-forward hop test was more strongly associated with outcomes than another, we performed sensitivity analysis using data from each individual hop test. Other functional tests (eg, side hop, vertical hop, one-leg rise, step-down test) were analysed separately as they measure different constructs of lower-limb function. We also assessed the relationship between a test battery (ie, including hop tests, and in some studies measures of strength or other biomechanical measures of function) and future knee-related outcomes for studies that combined several functional tests into an overall battery. We planned an a priori secondary analysis to separately assess subgroups (eg, male vs female, non-operative vs operative). Meta-regression was also used to assess the effect of follow-up time and proportion of females on the relationship between single-forward and repeated-forward hop tests and each outcome.

When two or more studies reported the results from the same hop test category (ie, single-forward hop or repeated-forward hop) and association to the same knee-related outcome,
meta-analysis was performed (Stata, V.17.0) if the OR and 95% CI could be calculated. The OR represents the likelihood of a higher LSI relating to a successful outcome. We used random-effects models as differences in population, functional test administration and outcome definition were expected. When ORs were not reported in the included studies, or could not be calculated from the number of individuals achieving a successful outcome, the standard mean difference (of functional performance between participants with and without a successful outcome) was calculated and then transformed to an OR; this was done using the method described in the Cochrane Handbook.29 Data from adjusted analyses were extracted wherever possible. An OR > 1 indicates a higher LSI being associated with successful future outcomes for return to sport rates, symptoms and function and success with ACL deficiency, while an OR < 1 indicates a higher LSI being associated with successful future outcomes for a subsequent knee injury and knee OA (ie, no further knee injury and no knee OA). Unplanned sensitivity analysis was completed to compare outcomes from The Knee injury and Osteoarthritis Outcome Score (KOOS; Englund symptomatic knee criteria)32 and International Knee Document Committee Subjective Knee Evaluation Form 2000: (1) scores below the 15th percentile of uninjured people33, (2) less than ‘Patient Acceptable Symptom State’ (PASS) threshold (ie, <75.9)34. Success with ACL deficiency (successful with non-operative treatment) — Instability or giving way episodes or Subsequent ACLR Subsequent knee injury after ACLR — Injury to the ipsilateral or contralateral knee Knee osteoarthritis — Presence of structural features on imaging which reach an established expert or consensus threshold of magnitude and characteristics to be termed osteoarthritis (eg, Kellgren and Lawrence grade). 

### Risk of bias and certainty of evidence
Our risk of bias assessment had excellent interrater reliability between the two independent assessors (k=97%). Only four studies conducted power analysis, 10 had acceptable loss to follow-up (mean lost to follow-up for all included studies was 19%) and 16 adjusted for potential confounders (online supplemental appendix 5). The overall certainty of evidence for all estimates was rated as very low using the GRADE tool. Evidence was downgraded based on study limitations, study design (observational studies), risk of bias, inconsistency, indirectness and publication bias (online supplemental appendices 6, 7).

### Synthesis of results for limb symmetry as a continuous score
**Return to sport**
Meta-analysis of nine studies (1020 participants) revealed that higher single-forward and repeated-forward hop test scores reported as an LSI up to 1 year following ACLR were associated with higher return to sport rates 1.5–3 years post-ACLR (OR 2.15; 95% CI 1.30 to 3.54; I² 70.6% and OR 2.11; 95% CI 1.23 to 3.60; I² 72.7%, respectively) (figures 2 and 3). One study not included in meta-analysis reported no association between single-forward hop test scores 3 years post-ACLR injury or reconstruction, and level of sports participation at 5 years.15

**Self-reported symptoms and function**
Meta-analysis of 10 studies (3107 participants) indicated that higher LSI scores on the single-forward and repeated-forward hop tests up to 4 years post-ACLR injury were associated with better self-reported knee symptoms and function 1–37 years following injury (see online supplemental appendix 3 for study outcome definitions). All but two studies38 39 reported functional test results as an LSI, of which one38 could not be estimated and was reported separately. All studies included the single-forward hop and 32 studies (78%) included a repeated forward hop (crossover hop; n=22, triple hop; n=17; 6m-timed hop; n=15) (online supplemental appendix 4). It was not feasible to analyse subgroups due to insufficient data for stratification.

### Table 1 Definitions of dichotomised knee-related outcome

<table>
<thead>
<tr>
<th>Knee-related outcome</th>
<th>Definition of unsuccessful outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return to sport</td>
<td>Not returning to previous level of sport or Not maintaining level of sport (after returning to sport)</td>
</tr>
<tr>
<td>Self-reported symptoms and function</td>
<td>The Knee injury and Osteoarthritis Outcome Score ‘Englund symptomatic knee criteria’ (Englund et al)32: QOL subscale ≤87.5 and at least two other subscales below following thresholds: pain ≤86.1, symptoms ≤85.7, AD L≤86.8, sport/rec ≤85.0. International Knee Document Committee Subjective Knee Evaluation Form 2000: (1) scores below the 15th percentile of uninjured people33, (2) less than ‘Patient Acceptable Symptom State’ (PASS) threshold (ie, &lt;75.9)34</td>
</tr>
<tr>
<td>Success with ACL deficiency (successful with non-operative treatment)</td>
<td>Instability or giving way episodes or Subsequent ACLR</td>
</tr>
<tr>
<td>Subsequent knee injury after ACLR</td>
<td>Injury to the ipsilateral or contralateral knee</td>
</tr>
<tr>
<td>Knee osteoarthritis</td>
<td>Presence of structural features on imaging which reach an established expert or consensus threshold of magnitude and characteristics to be termed osteoarthritis (eg, Kellgren and Lawrence grade)</td>
</tr>
</tbody>
</table>

ACL, anterior cruciate ligament; ACLR, anterior cruciate ligament reconstruction; ADL, activities of daily living; QOL, quality of life.

### RESULTS
**Study characteristics**
From an initial yield of 10272 studies, 42 were included in this review consisting of a total 13150 adolescents and adults (44% female) (figure 1, table 2). Of the 42 studies, 38 (95%) assessed functional performance up to 1 year after ACL reconstruction (ACLR) (hop tests conducted at 19 months in one study,35 3 years in another13 and 4 years post-ACLR injury in two studies16 37 with follow-up assessment of outcomes ranging between 1 and 37 years following injury (see online supplemental appendix 3 for study outcome definitions). All but two studies38 39 reported functional test results as an LSI, of which one38 could not be estimated and was reported separately. All studies included the single-forward hop and 32 studies (78%) included a repeated forward hop (crossover hop; n=22, triple hop; n=17; 6m-timed hop; n=15) (online supplemental appendix 4). It was not feasible to analyse subgroups due to insufficient data for stratification.
Review

Knee osteoarthritis

Meta-analysis of four studies (444 participants) revealed no association between single-forward and repeated-forward hop tests and developing knee OA between 1 and 37 years after ACL injury (OR 0.60; 95% CI 0.25 to 1.43; I^2 75.8% and OR 0.39; 95% CI 0.13 to 1.20; I^2 76.7%, respectively) (figures 2 and 3). Sensitivity analysis indicated that higher scores on the crossover hop (but not other repeated-forward hop tests) may be prognostic for structural knee OA (OR 0.14 95% CI 0.05 to 0.34) (online supplemental appendix 10).

Other outcomes

Single studies reported the association between functional tests and outcomes of knee biomechanics and self-efficacy. Higher scores on a single-leg step-down test at 3 months post-ACLR significantly correlated with greater knee flexion excursion (r=0.65, p=0.001) and knee extensor moment (r=0.54, p=0.001) during running 6 months post-ACLR. LSI for the single-forward hop, within the first year after ACL injury, was moderately correlated with worse knee-related self-efficacy after 1 year of rehabilitation alone or ACLR (r=0.31, p=0.001).

Synthesis of results for dichotomised limb symmetry scores

Meta-analyses of studies using a dichotomous threshold of ≥90% limb symmetry demonstrated consistent findings compared with using continuous LSI measures for outcomes of...
Table 2  Characteristics of included studies grouped by outcome

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Participants*†</th>
<th>Female‡ (%): Time of functional test§</th>
<th>Ages±SD, years**</th>
<th>BMI±SD, kg/m²**</th>
<th>Functional tests</th>
<th>Mean follow-up¶</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return to sport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ardern et al, 2015††† Australia</td>
<td>Prospective ACLR: 94</td>
<td>38%</td>
<td>1 year</td>
<td>29±8</td>
<td>NR</td>
<td>Single hop</td>
<td>Triple hop</td>
</tr>
<tr>
<td>Ebert and Annear, 2019†††† Norway</td>
<td>Prospective ACLR: 50</td>
<td>36%</td>
<td>1 year</td>
<td>26±10</td>
<td>24.8±4.0</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Falde et al, 2021†† Norway</td>
<td>Prospective ACLR: 103</td>
<td>47%</td>
<td>9 months</td>
<td>29±10</td>
<td>NR</td>
<td>Test battery only§§</td>
<td></td>
</tr>
<tr>
<td>Kinguchi et al, 2020 Japan</td>
<td>Prospective ACLR: 124</td>
<td>59%</td>
<td>6 months</td>
<td>15±2¶¶</td>
<td>NR</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Molander and Rosberg, 2009†††† Norway</td>
<td>Prospective ACLR: 52 ACLR: 50</td>
<td>45%</td>
<td>6 months</td>
<td>28±9</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Nawasreh et al, 2017††† USA</td>
<td>Prospective ACLR: 95</td>
<td>34%</td>
<td>6 months</td>
<td>27±10</td>
<td>24.9±3.6</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Tolle et al, 2017 USA††</td>
<td>Prospective ACLR: 115</td>
<td>77%</td>
<td>8 months</td>
<td>17±2</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Webster et al, 2019 Australia†††</td>
<td>Prospective ACLR: 222</td>
<td>40%</td>
<td>1 year</td>
<td>26±9</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Welling et al, 2020††† Netherlands</td>
<td>Prospective ACLR: 64</td>
<td>31%</td>
<td>10 months</td>
<td>27±8</td>
<td>NR</td>
<td>Single hop</td>
<td>Triple hop</td>
</tr>
<tr>
<td>Symptoms and function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cristiani et al, 2020†† Sweden</td>
<td>Prospective ACLR: 2335</td>
<td>49%</td>
<td>6 months</td>
<td>29±10</td>
<td>NR</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Culver et al, 2016†††† Australia</td>
<td>Prospective ACLR: 93</td>
<td>40%</td>
<td>1 year</td>
<td>27 (13) (median and IQR)</td>
<td>NR</td>
<td>Single hop</td>
<td>Triple hop</td>
</tr>
<tr>
<td>Ericsson et al, 2013†††† Denmark</td>
<td>Prospective ACLD: 42 Delayed ACLR: 20 Early ACLR: 45</td>
<td>26%</td>
<td>37 weeks (median)</td>
<td>26±5</td>
<td>23.9±2.6</td>
<td>Single hop</td>
<td>Square hop</td>
</tr>
<tr>
<td>Filbay et al, 2021††† Sweden</td>
<td>Prospective ACLR or ACLD: 109</td>
<td>28%</td>
<td>4 years</td>
<td>24±6†††</td>
<td>27 (24–29) (median and IQR)</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Løgterstedt et al, 2012†††† USA</td>
<td>Prospective ACLR: 85</td>
<td>45%</td>
<td>6 months</td>
<td>26 (15–54) (median and range)</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Månsson et al, 2013††† Sweden</td>
<td>Prospective ACLR: 73</td>
<td>30%</td>
<td>Pre-ACLR</td>
<td>24 (14–40) (median and range)</td>
<td>NR</td>
<td>Single hop</td>
<td>Triple hop</td>
</tr>
<tr>
<td>McGraith et al, 2017††† Australia</td>
<td>Prospective ACLR: 64</td>
<td>31%</td>
<td>24 weeks</td>
<td>28</td>
<td>25</td>
<td>Single hop</td>
<td>6m-timed hop</td>
</tr>
<tr>
<td>Øiestad et al, 2012†††† Norway</td>
<td>Prospective ACLR: 181</td>
<td>42%</td>
<td>1 year</td>
<td>27±9</td>
<td>NR</td>
<td>Single hop</td>
<td>Triple hop</td>
</tr>
<tr>
<td>Stropnik et al, 2020†† Slovenia</td>
<td>Prospective ACLR: 60</td>
<td>31%</td>
<td>Pre-ACLR</td>
<td>32±11§§§</td>
<td>25.7±3.5§§§</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Success with ACL deficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Button et al, 2006†††† Wales</td>
<td>Prospective ACLD: 42¶¶¶</td>
<td>40%</td>
<td>5 months</td>
<td>28±7</td>
<td>NR</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Elten et al, 2010†† Norway/USA</td>
<td>Prospective ACLR: 145</td>
<td>52%</td>
<td>60 days</td>
<td>26 (14–47) (median and range)</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Ekds et al, 2019†† Norway</td>
<td>Prospective ACLR: 44</td>
<td>34%</td>
<td>2 years</td>
<td>11±1</td>
<td>24.7 (range: 16.6–40.8)</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Fitzgerald et al, 2000†† USA</td>
<td>Prospective ACLR: 28</td>
<td>25%****</td>
<td>4 weeks (median)</td>
<td>29±11</td>
<td></td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
<tr>
<td>Grindem et al, 2018††† USA</td>
<td>Prospective ACLR: 118</td>
<td>51%</td>
<td>2 months</td>
<td>28±10</td>
<td>24.2±3.8</td>
<td>Single hop</td>
<td>Cross-over hop</td>
</tr>
<tr>
<td>Subsequent injury after ACLR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beischer et al, 2020†† Sweden</td>
<td>Prospective ACLR: 159</td>
<td>44%</td>
<td>9 months</td>
<td>20±3</td>
<td>NR</td>
<td>Test battery only§§</td>
<td></td>
</tr>
<tr>
<td>Cristiani et al, 2021††† Sweden</td>
<td>Prospective ACLR: 6510</td>
<td>44%</td>
<td>6 months</td>
<td>28±10</td>
<td>24.2±3.4</td>
<td>Single hop</td>
<td></td>
</tr>
<tr>
<td>Fältström et al, 2021†† Sweden</td>
<td>Prospective ACLR: 117</td>
<td>100%</td>
<td>19 months</td>
<td>20±2</td>
<td>NR</td>
<td>Single hop</td>
<td>Crossover hop</td>
</tr>
</tbody>
</table>

Continued
Table 2 Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Participants*†</th>
<th>Female‡ (%)</th>
<th>Time of functional tests††</th>
<th>Age±SD, years**</th>
<th>BMI±SD, kg/m²***</th>
<th>Functional tests</th>
<th>Mean follow-up¶</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grindem et al, 2016†††† Norway/USA</td>
<td>Prospective</td>
<td>ACLR: 106</td>
<td>51%</td>
<td>6 months†††</td>
<td>25±7</td>
<td>NR</td>
<td>Single hop</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>King et al, 2021†††† Ireland</td>
<td>Prospective</td>
<td>ACLR: 115</td>
<td>0%</td>
<td>9 months</td>
<td>21±4</td>
<td>NR</td>
<td>Single hop</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Kyntis et al, 2016†††† Qatar</td>
<td>Prospective</td>
<td>ACLR: 158</td>
<td>0%</td>
<td>247 days</td>
<td>21±4</td>
<td>NR</td>
<td>Single hop</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Nawasreh et al, 2016†† USA</td>
<td>Prospective</td>
<td>ACLR: 95</td>
<td>32%</td>
<td>6 months</td>
<td>27±10</td>
<td>25±3.8</td>
<td>Test battery only§§</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Sousa et al, 2017††† USA</td>
<td>Prospective</td>
<td>ACLR: 223</td>
<td>59%</td>
<td>6 months</td>
<td>26±11</td>
<td>25.8±4.5</td>
<td>Test battery only§§</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>van Melick et al, 2021†††† Netherlands</td>
<td>Prospective</td>
<td>ACLR: 144</td>
<td>31%</td>
<td>1 year</td>
<td>24±7</td>
<td>NR</td>
<td>Single hop</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Webster et al, 2019†† Australia</td>
<td>Prospective</td>
<td>ACLR: 329</td>
<td>39%</td>
<td>1 year</td>
<td>17±2</td>
<td>NR</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Wellbandt et al, 2017††† USA</td>
<td>Prospective</td>
<td>ACLR: 70</td>
<td>33%</td>
<td>6 months</td>
<td>26±10</td>
<td>24.9±3.8</td>
<td>Single hop</td>
<td>2 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Kline et al, 2007†††† Australia</td>
<td>Prospective</td>
<td>ACLR: 106</td>
<td>51%</td>
<td>6 months†††</td>
<td>25±7±11</td>
<td>NR</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Filbay et al, 2021†††† Sweden</td>
<td>Prospective</td>
<td>ACLR or ACLD: 127</td>
<td>28%</td>
<td>4 years</td>
<td>23 (19–28) (median and IQR)†††</td>
<td>27 (24–28) (median and IQR)§§§</td>
<td>Single hop</td>
<td>32–37 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Janssen et al, 2013†††† Netherlands</td>
<td>Prospective</td>
<td>ACLR: 86</td>
<td>34%</td>
<td>Pre-ACLR</td>
<td>31±8§§§</td>
<td>24.5±3.7§§§</td>
<td>Single hop</td>
<td>10 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Patterson et al, 2018†††† Australia</td>
<td>Prospective</td>
<td>ACLR: 78</td>
<td>38%</td>
<td>1 year</td>
<td>28±14</td>
<td>25.7±4.2</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Przeczewski et al, 2008††††† Australia</td>
<td>Prospective</td>
<td>ACLR: 149</td>
<td>47%****</td>
<td>1 year</td>
<td>25 (13–52) (median and range)****</td>
<td>NR</td>
<td>Single hop</td>
<td>10 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Wellbandt et al, 2018†† USA</td>
<td>Prospective</td>
<td>ACLR or ACLD: 58</td>
<td>35%*****</td>
<td>6 months</td>
<td>28±11</td>
<td>NR</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Other (studies not included in meta-analysis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Flosadottir et al, 2016†† Sweden</td>
<td>Prospective</td>
<td>ACLR or ACLD: 54</td>
<td>28%</td>
<td>3 years</td>
<td>29±5</td>
<td>24.6±3.4</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Flosadottir et al, 2016†††† Sweden (Knee self-efficacy)</td>
<td>Prospective</td>
<td>ACLR: 20</td>
<td>28%</td>
<td>37 weeks (median)</td>
<td>25±5</td>
<td>23.8±2.8</td>
<td>Single hop</td>
<td>5 years</td>
<td>(online supplemental appendix 11).</td>
</tr>
<tr>
<td>Kline et al, 2005†††† USA (Knee biomechanics)</td>
<td>Prospective</td>
<td>ACLR: 30</td>
<td>53%</td>
<td>3 months</td>
<td>21±7</td>
<td>NR</td>
<td>Single leg step down</td>
<td>6 months</td>
<td>(online supplemental appendix 11).</td>
</tr>
</tbody>
</table>

*As listed in the original paper (may differ with extra data provided by the author).†Studies were of mixed sporting populations apart from those indicated otherwise.‡Most studies reported on percentage of male v female sex though used female/women and male/men terminology interchangeably.¶Mean, unless indicated otherwise.††Time or follow-up was number of years from the time of injury (deficient cohorts) or ACLR (reconstructed cohorts) unless indicated otherwise.**Assessed at time of functional assessment unless indicated otherwise.†††Extra data provided by authors to enable synthesis of results and dichotomous outcome comparison.‡‡Hop test results also compared to self-reported symptoms and function.§§Only results from test battery able to be included for analysis.¶¶Assessed at time of follow-up.***Hop test results also compared to success with ACL deficiency.††††Assessed at time of injury.§§§§Hop test results also compared with knee osteoarthritis.****Assessed at time of primary ACLR.†††††Functional testing was taken at 6 and 12 months post-ACLR.‡‡‡‡Studies included professional male athletes (all other studies were of a general population).§§§§Outcome used was knee self-efficacy.****Outcome used was biomechanical variables. ACLD, anterior cruciate ligament deficient; ACLR, anterior cruciate ligament reconstruction; BMI, body mass index; NR, not reported.

**Synthesis of results for a test battery**

Passing a test battery (ie, achieving LSI ≥90% on all tests) was associated with greater odds of returning to sport (OR 3.33; 95% CI 1.05 to 10.53; I² 55.3%) (online supplemental appendix 14). Passing a battery of tests was not associated with reduced odds of subsequent knee injury after ACLR (OR 0.62; 95% CI 0.26 to 1.48; I² 67.3%) (online supplemental appendix 14). Sensitivity analysis revealed that passing test batteries made up exclusively of hop tests was associated with avoiding subsequent knee injury (OR 0.29, 95% CI 0.10 to 0.85; I² 44.3%). Conversely, passing test batteries that included strength measures along with hops tests was not associated with avoiding return to sport (online supplemental appendices 11, 12). Effect sizes for self-reported symptoms and function and success with ACL deficiency outcomes were attenuated where dichotomous LSI data were used (online supplemental appendices 11, 12). Meta-analysis of two studies also found no association with side hop 2 years post-ACLR and subsequent knee injury (online supplemental appendix 13).
Figure 2 Forest plot displaying the association between single-forward hop results (continuous Limb Symmetry Index scores) and future outcomes. Subsequent knee injury (1.31; 95%CI 0.63 to 2.71 I^2 22.4%) (online supplemental appendix 15).

Sensitivity analysis of associations found between hop test and knee-related outcomes
Sensitivity analysis separating each repeated-forward hop test demonstrated that triple hop, crossover hop and 6m-timed hop had similar effect sizes for each outcome, with no significant between-test heterogeneity observed: RTS (I^2 71.8%, p=0.645), self-reported symptoms and function (I^2 51.9%, p=0.978), success with ACL deficiency (I^2 19.6%, p=0.843) and subsequent injury after ACLR (I^2 1.2%, p=0.175). Between test heterogeneity was significantly different with individual repeated-forward hops tests and knee OA, with the crossover hop demonstrating a significantly larger association with knee OA than the 6m-timed hop or triple hop test (I^2 76.7%, p=0.014) (online supplemental appendix 10). Meta-regression indicated that follow-up time and proportion of females did not influence the relationship between single-forward and repeated-forward hop tests and all outcomes (online supplemental appendix 16).

DISCUSSION
This systematic review of 42 studies and 13 150 adolescents and adults following ACL injury, found an increased odds of poor knee-related outcomes after 1–37 follow-up years in individuals not achieving adequate LSI on unilateral functional tests. These findings highlight the need to optimise knee function in the first 1–2 years post-ACLR injury and reconstruction to reduce the long-term burden. Caution is warranted when interpreting these findings given the very low certainty of evidence associated with the results. The low certainty of evidence points to the need for high quality longitudinal data that is well controlled for confounders existing (and/or changing) during the long duration between exposure (hop tests) and outcomes (eg, physical activity, muscle strength).

Figure 3 Forest plot displaying the association between repeated-forward hop results (continuous Limb Symmetry Index scores) and future outcomes. In brackets are the specific repeated-forward hop used in analysis.

What is the prognostic value of hop tests for return to sport, reinjury and self-reported outcomes?
Achieving a higher LSI on the single-forward and repeated-forward hop tests was associated with twice the odds of future return to sport (OR 2.15; 95%CI 1.30 to 3.54; I^2 70.6% and OR 2.10; 95%CI 1.23 to 3.60; I^2 72.7%, respectively) (figures 2 and 3) with 9/10 studies in this analysis using pre-injury level of sport as the measure of successful return. While this indicates the potential importance of achieving adequate functional performance to facilitate a return to preinjury sport, those who return to high-impact and pivoting sports are known to be at higher risk of reinjury.25 Our pooled results suggest a slightly elevated risk of reinjury in those with a higher hop score, which may reflect higher return to sport rates in those with better hop performance. There is much debate regarding the importance of achieving common return to sport criteria (ie, LSI ≥90% on a functional test battery including hop tests) to reduce reinjury risk.20 42 43 For example, some original studies and systematic reviews report that meeting functional criteria increases reinjury risk,20 42 while others report a reduced risk of reinjury on passing return to sport criteria.21 25 Inconsistencies are likely explained by whether sport participation (both level and type of sport) exposure was accounted for and whether some participants who may initially fail test batteries continued rehabilitation until passing.42 43 Individuals following ACL injury and reconstruction should be educated about the risk of reinjury when returning to sport regardless of adequate functional performance and consider other factors known to increase re-injury risk (eg, time postsurgery, younger age, symmetry of quadriceps strength, fear).25 44 45

We also observed that higher LSI on single-forward and repeated-forward hop tests was associated with more than twofold and fourfold higher odds of better self-reported symptoms and function, respectively (OR 2.51; 95%CI 1.62 to
Can hop tests tell us who is going to get post-traumatic structural OA?

Knee OA is a burdensome sequela of traumatic knee injury with its risk not reduced with reconstruction.47 We found achieving LSI ≥90% on the single-forward hop was associated with reduced odds of developing structural knee OA (LSI ≥90%; OR 0.46; 95%CI 0.23 to 0.94; F 54.5%) (online supplemental appendix 11), though this was attenuated when analysing hop test scores as continuous variables. The stronger association with OA when using a dichotomous hop score, may be due to including different studies in each meta-analysis (n=2 not included in both meta-analyses). In sensitivity analysis data from a single study48 indicated that higher scores on the crossover hop may be prognostic for structural knee OA (OR 0.14; 95%CI 0.05 to 0.34) (online supplemental appendix 10). Structural knee OA was defined differently across included studies (eg, joint space narrowing, Kellgren and Lawrence classification and adding symptoms to definition) and there was no observable trend for the impact of hop test performance and structural OA based on length of follow-up. In the context of OA prevention, addressing impairments in hop tests (ie, quadriceps weakness) may provide a target for intervention though RCTs are required to confirm this.48 49

Is one test better than another or is a battery of tests needed?

Generally, results were similar between single- and repeated-forward hop tests for most outcomes. As studies often included multiple tests, we created a hierarchy for reporting repeated-forward hop tests (ie, crossover hop, triple hop, 6m-timed hop). Despite tests measuring somewhat different constructs (ie, crossover hop includes lateral and forward movement and 6m-timed hop measuring speed of completion) each individual repeated-forward hop tests found similar associations (online supplemental appendix 10). Current clinical trial evidence suggests that trialling rehabilitation without surgery is not inferior to having early ACLR.50 In this context, our results highlight the importance of conducting hop tests during the initial rehabilitation period to monitor function, which might help to determine an individual’s ability to cope with an ACL deficient knee, alongside other factors such as desired level of activity (eg, return to cutting sports).

A battery of tests is recommended to determine readiness to return to sport, although it is difficult to compare batteries given the inconsistency in tests included across different studies.51 The common definition for ‘passing’ a test battery is LSI ≥90% on all tests included.51 52 53 Results from our review suggest that both single-forward and repeated-forward hops are similar to a test battery as prognostic indicators of return to sport. Passing a test battery (consisting only of hop tests) was associated with reduced odds of further knee injury more than threefold compared with not passing (OR 0.29; 95% CI 0.10 to 0.85) (online supplemental appendix 15). In contrast, we found that a higher LSI on individual hop tests (figures 2 and 3) or passing test batteries (including those with hop tests and strength measures) was not associated with subsequent knee injury (online supplemental appendix 14). There is no agreement across our analyses reflective of findings from various recent systematic reviews.46 20–22

Compared with these reviews, we included different studies and we also pooled knee injuries under one outcome (eg, ipsilateral and contralateral reinjury, other knee injury). Given our results and those from previous reviews, we suggest that a test battery is likely needed to identify those at higher risk of future knee injury. Unfortunately for clinicians, exactly how many and what particular combination of tests to use remains unclear but current guidelines recommend using test batteries that include hop tests and measures of strength.53 54

The prognostic capacity of less common functional tests

Less common tests of functional performance included in this review were the one-leg rise, side hop and vertical hop. We found that performance on the one-leg rise was associated with future self-reported symptoms and function. The one-leg rise test may be a surrogate of quadriceps (and lower-limb) strength55 and is becoming more popular in clinical settings due to the lack of equipment required compared with isokinetic dynamometry. Higher side hop performance was not associated with subsequent knee injury after ACLR (online supplemental appendix 13). Higher LSI on the vertical hop was associated with success with an ACL deficient knee56 and may be a better alternative to forward hop tests to assess knee function because the knee contributes more to vertical height than horizontal distance (which is more a function of hip/ankle power).54–56

In our review, we only compared the quantitative achievement of a functional test and results do not account for movement quality or other aspects of how tests were completed. Sixty-five per cent of included studies were deemed high risk of bias in terms of how the hop tests were assessed or completed, identifying the need for greater consistency to strengthen comparisons. It is also becoming more apparent that poor movement quality (eg, measured using two-dimensional or three-dimensional kinematics) is another important construct to assess given its association to poor outcome despite restoration of quantitative function (eg, LSI ≥90%).57 The importance of qualitative assessment also highlights the limitation of the LSI as a measure of recovered function. The contralateral (reference) limb may deteriorate after ACL injury with disuse and central inhibition58 59 and thus not be an accurate preinjury reference standard. For this reason, other criteria, such as an estimated preinjury capacity (ie, contralateral hop test result immediately after ACL injury) have been proposed as a more suitable measure of preinjury function as well as normative values from well-matched controls.51 58

Limitations

Results from our review are limited by the very low certainty of evidence. The level of evidence (GRADE assessment) required downgrading due to observational studies and large I² values. These values indicated heterogeneity among pooled studies that was largely unexplained as our sensitivity analyses were unable to determine a source of this. Most studies had unacceptable lost to follow-up (eg, >15%), which may indicate attrition bias that may have altered results. Self-reported symptoms and function at follow-up may reflect higher scores at baseline rather than improvement over time, although other studies have shown only fair cross-sectional correlation between functional tests and self-reported symptoms and function.46 Nevertheless, our results demonstrate the prognostic value of these functional tests, which may aid clinicians in identifying those at higher risk of long-term burden.
skewing results. While eight of the 42 included studies accounted for baseline differences (eg, age, sex, body mass index), the majority did not account for these factors and their potential impact on results. Indirectness was also found with time between test and follow-up differing across studies and though this enhances generalisability, it also reduces confidence in the estimated effects. Graft type used for ACLR was also not able to be included as part of our analysis to compare surgery types as the data were not available. Finally, we were unable to provide an estimate of the relationship between functional test results and future psychological recovery (eg, confidence, fear) as no studies had available data evaluating this. However, we recognise the growing awareness of the importance of restoring confidence and reducing fear after ACL injury\(^{19,20}\) and this should be investigated further.

**CONCLUSION**

In this systematic review, we found unilateral tests of lower-limb function can be prognostic for future knee-related outcomes in individuals after ACL injury, though with very low certainty evidence. Early in rehabilitation, hop testing may be used as part of criteria to consider non-operative management and along with other tests may provide insight into potential future symptoms and function over the long term, including the development of post-traumatic knee OA. Hop tests are also recommended to be used as part of return to sport testing, though achieving adequate performance may not reduce subsequent knee injuries.

**Twitter** Thomas J West @tswennot, Andrea M Bruder @AndreaBruder, Kay M Crossley @KayCrossley and Adam G Culvenor @agculvenor

**Contributors** TJW, AGC, AMB and KMC designed the study. TJW and AGC completed all searches, study selection (including inclusion and exclusion of abstracts), TJW and AGC completed all data extraction. TJW and AGC and KMC planned the analyses, TJW completed the meta-analyses and all authors interpreted the data. TJW wrote the initial draft and all authors critically revised the manuscript for important intellectual content approved the final version of the manuscript.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**ORCID iDs**

Thomas J West http://orcid.org/0000-0002-6297-1094

Andrea M Bruder http://orcid.org/0000-0001-5422-5756

Kay M Crossley http://orcid.org/0000-0001-5892-129X

Adam G Culvenor http://orcid.org/0000-0001-9491-0264

**REFERENCES**


8. Davies WT, Myer GD, Read PI. Is it time we better understood the tests we are using for return to sport decision making following ACL reconstruction? A critical review of the hop tests. *Sports Med* 2020;50:485–95.


Review


58 Ebert JR, Annear PT. ACL reconstruction using autologous Hamstrings augmented with the lip arthroplasty and reconstruction system provides good clinical scores, high levels of satisfaction and return to sport, and a low Reteat rate at 2 years. Orthop J Sports Med 2019;7:2325967119879079–79.


