Comparative Effectiveness and Safety of Allografts and Autografts in Posterior Cruciate Ligament Reconstruction Surgery: A Systematic Review



Cecilia de Villiers, M.Sc., B.Sc., Gregor Goetz, M.S.Sc., M.P.H., Patrick Sadoghi, M.D., Ph.D., M.B.A., and Sabine Geiger-Gritsch, D.M.Sc., M.H.Sc., M.Pharm

Purpose: To evaluate the evidence of the effectiveness and safety of allografts compared to autografts in posterior cruciate ligament reconstruction. Methods: Four electronic databases were systematically searched for eligible randomized controlled studies. Crucial effectiveness outcomes included patient-reported function, activity level and symptoms, clinical knee stability, health-related quality of life, and patient satisfaction. Safety was evaluated through graft failures, revisions, reruptures and complications. The internal validity of the studies was assessed by the Cochrane risk of bias tool, and the strength of the evidence was judged according to the Grading of Recommendations, Assessment, Development and Evaluation (GRADE). Results: Two randomized controlled studies were included: 50 patients were analyzed in the allograft group and 58 in the autograft group. No statistically significant postoperative differences were reported between the groups for patient-reported function, activity levels or symptoms. One study reported a statistically significant difference in instrumented postoperative anteroposterior knee laxity favoring stability in autografts. This difference is, however, not relevant in the clinical setting. Insufficient evidence was found to judge safety outcomes and because complications were poorly measured, and none of the studies reported on graft failure, revision or rerupture rates. The studies were judged with unclear to high risk of bias. The strength of the evidence for effectiveness and safety was judged to be low to very low, according to GRADE. Conclusions: Allografts may be comparable to autografts for crucial effectiveness outcomes, but insufficient evidence was found to judge crucial safety outcomes due to poor reporting of safety measures and outcomes. Results should be interpreted with caution because there is lack of good-quality evidence to support the superiority of allografts over autografts due to the high risk of bias in the primary studies and overall very low strength of the body of evidence according to GRADE. Level of Evidence: Systematic review of Level II studies.

A in annual incidence rate of 2 cases per 100,000 individuals is reported for isolated posterior cruciate ligament ruptures. Historically, a conservative

From the HTA Austria - Austrian Institute for Health Technology Assessment GmbH (Former: Ludwig Boltzmann Institute for Health Technology Assessment), Vienna, Austria (C.d.V., G.G., S.G.-G.); and Department of Orthopaedics and Trauma, Medical University of Graz, Graz, Austria (P.S.).

The authors report that they have no conflicts of interest in the authorship and publication of this article. Full ICMJE author disclosure forms are available for this article online, as supplementary material.

Received December 11, 2019; accepted July 30, 2020.

Address correspondence to Cecilia de Villiers, M.Sc., B.Sc., HTA Austria - Austrian Institute for Health Technology Assessment GmbH (Former: Ludwig Boltzmann Institute for Health Technology Assessment), Garnisongasse 7/Top 20, 1090, Vienna, Austria. E-mail: cila.devilliers@gmail.com

Published by Elsevier on behalf of the Arthroscopy Association of North America. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

2666-061X/191489

https://doi.org/10.1016/j.asmr.2020.07.017

management approach has been followed and because of the intrinsic healing capability of the PCL, surgical intervention has been reserved mainly in the event of failed conservative management or for the treatment of associated injuries.^{2,3} PCL reconstruction as first-line routine care is still debated⁴ because surgical intervention is not without its own challenges; there is a reported graft failure rate of 11.6% (range of 2.3%-30%)⁵ and a revision rate of 12.6% in a highly active military population.⁶ Complications such as rerupture or graft failures due to technical errors, including poor graft size, inadequate graft strength and improper placement of the tibial or femoral tunnels, ^{7,8} have been reported. Significant improvements were observed in post-PCL reconstruction functional scores and knee stability, although this seems to be valid only for moderate levels of activity and a fair return to sport level. A careful consideration of surgical-intervention decisions should, therefore, include multiple factors,

such as the patient's level of activity, the functional demands on the knee, the surgeon's experience, and the presentation of other associated meniscal or ligament injuries.²

The graft type to be used is a further decision to consider because autografts and allografts are potential sources of PCL reconstructions. The chosen graft should resemble the original PCL anatomy, taking into consideration that each graft is associated with different benefits and limitations. Autografts do not generally pose the risk of rejection or disease transmission, although they are associated with longer surgical times, postoperative donor-site pain and morbidity and weakness of the donor extensor or flexor muscle groups. 10 Allografts provide an alternative, but they are associated with their own limitations, including potential graft rejection, disease transmission¹¹ and weakening of the graft structure as a result of the sterilization process.^{2,12} This increases the potential risk of delayed healing, remodeling and integration of the graft; tunnel widening, cyst formation and effusions have been reported. Limited graft availability and the associated costs are also factors to consider with the choice of allograft.²

The current literature concerning systematic reviews reports postoperative improvement of clinical and patient-reported outcomes following PCL constructions in both allograft and autograft groups. 12,14,15 Controversy remains in regard to whether 1 graft type provides better outcomes and is superior to the other in PCL reconstruction. 16 The purpose of this study was to evaluate the effectiveness and safety evidence of allografts compared to autografts in PCL reconstructions. The authors evaluated the hypothesis that the effectiveness and safety outcomes of allografts are superior compared to those of autografts in PCL reconstruction.

Methods

This systematic review was based on the European Network of Health Technology Assessment (EUnetHTA) Core Model domains (version 3.0)¹⁷ and was reported via the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The project was commissioned by the Austrian Ministry of Health as part of the decision support series for Health Technology Assessments.

The systematic literature and hand searches described below were used to identify relevant studies according to the Population, Intervention, Comparator, Outcomes and Study Design (PICOS).²⁰ Randomized controlled trial (RCT) study designs are deemed the best evidence for evaluating effectiveness and safety in Health Technology Assessments reports because of their robust internal validity and low risk of bias.²¹ The original systematic literature search was performed in 2018. A

recent updated manual search in the PubMed database as well as the Arthroscopy.com search function did not identify any additional eligible studies for inclusion between 2019 and 2020.

Literature Search Strategy

The Medline via Ovid, Embase, the Cochrane Library and Centre for Reviews and Dissemination electronic databases were systematically searched for eligible studies published from inception to December 2018. Eligible clinical trials and articles published in English or German were identified, and no other screening criteria were applied. A preceding hand search in PubMed identified published systematic reviews and metaanalyses, and their reference lists were used to identify additional relevant RCTs. All of these primary studies were also captured in the literature search, and any duplications were removed.²⁰ Three clinical trials registers were also searched. The original search was broad and included studies of the anterior cruciate ligament according to the commissioned project. Only the PCL reconstruction studies were selected as per the PRISMA diagram in Fig 1.

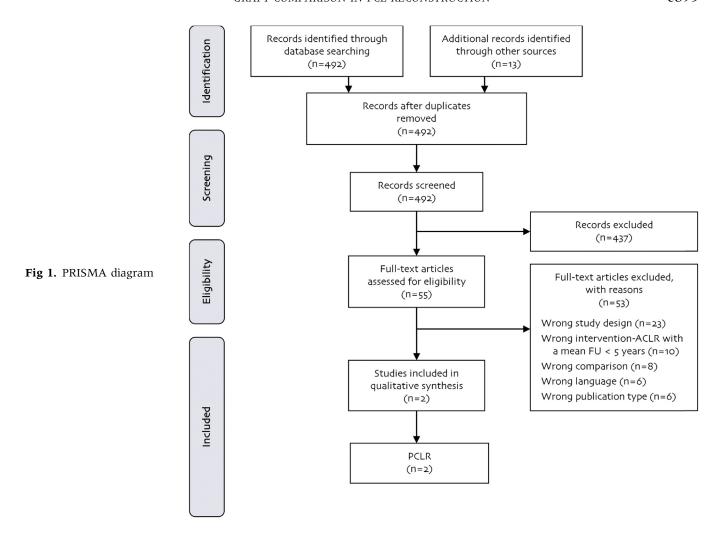
Study Eligibility, Inclusion and Exclusion Criteria and Study Selection

The inclusion criteria were in accordance with the study PICOS.²⁰ Patients who were eligible for PCL reconstruction, revision PCL reconstruction or PCL reconstruction as a result of multiligament knee injuries were included and compared in terms of crucial effectiveness and safety outcomes to other techniques of PLC reconstruction, such as autograft, synthetic graft and conservative management. For the purpose of this study, the focus was on PCL reconstruction. Revision PCL reconstruction and PCL reconstruction as a result of multiligament knee injuries were excluded due to the lack of evidence found to either support or reject the superiority of allografts over autografts.²⁰

RCTs were the chosen study design, ²¹with prospective, comparative cohort studies to be considered if no RCTs were available. Three authors (GG, SGG, CDV) conducted the study selection, and 2 authors (GG, SGG) screened the abstracts. The full-text screening was conducted by 2 authors (GG, CDV), and the third author (SGG) was consulted for consensus in the event of any unresolved discrepancies.

Outcome Measures

The following crucial outcomes were chosen in accordance with the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines²² to evaluate the relative effectiveness of allografts compared to autografts in PCL reconstruction: patient-reported function, activity levels and symptoms measured using validated instruments, such as the



Lysholm score, the Tegner score and the subjective part of the International Knee Documentation Committee (IKDC) scores.

Clinical knee stability measured by a validated instrument, such as the KT-1000 arthrometer and the reverse Lachman and Pivot shift tests, and health-related quality of life and patient satisfaction, as measured by validated instruments, were included as important outcomes but were not crucial for the assessment.

The crucial outcomes included to evaluate safety were graft failures, reruptures, reoperations and revisions, complications, and procedure-related mortality. Graft failures were chosen because they could result in further or revision surgery; however, it was not defined in a standardized way by the included studies. Reruptures, reoperations and revisions are synonyms for graft failures and were extracted separately. Complications were broad and included surgical complications, such as infection, anesthesia complications and complications associated with cruciate ligament surgery, such as

postoperative loss of motion, arthrofibrosis and donor site morbidity.

Data Extraction

Pre- and postoperative data were systematically extracted from the included studies. A single extraction method was used by 1 author, and a second author controlled the extracted data. Disagreements were resolved through discussion or in consultation with a third author. Table 1 summarizes the level of evidence²³ of the 2 included studies. The extraction table (Table 2) was populated with variables according to the PICOS study. This included the author, the year of publication, the study design, the intervention and comparators, the surgical procedure with study inclusion and exclusion criteria, the rehabilitation post surgery, the number of patients, their ages, sex, mean follow-up, loss to follow-up, and other patients' characteristics. Crucial effectiveness outcomes included patient-reported function, activity levels and symptoms and clinical knee stability. Crucial safety outcomes

Table 1. Level of Evidence²³

Study	Level of Evidence	Investigators
Prospective randomized comparison of knee	Level II	Li, Jia (author)
stability and proprioception for posterior cruciate		Kong, Fanlong (author)
ligament reconstruction with autograft, hybrid		Gao, Xianda (author)
graft and gamma-irradiated allograft		Shen, Yong (author)
		Gao, Shijun (author)
Comparison of autogenous and allogenous	Level II	Wang, Ching-Jen (author)
posterior cruciate ligament reconstruction of the		Chan, Yi-Sheng (author)
knee		Weng, Lin-Hsiu (author)
		Yaun, Li-Jen (author)
		Chen, Han-Shiang (author)

included graft failure rates, reruptures, revisions, overall complications, and other complications.

Study Quality Appraisal

The Cochrane Risk of Bias Tool²⁴ was used to assess the internal validity of the individual RCT studies. Two authors assessed the risk of bias independently, and any discrepancies were resolved through discussion. A third author was consulted if no consensus was reached.

Only postoperative data from the extraction tables were synthesized according to the GRADE²² methodology. Each outcome was individually judged according to the study design, risk of bias, inconsistency, indirectness, imprecision, and other considerations. The preoperative scores were excluded from the synthesis because the primary studies reported mainly the postoperative differences for the comparison between the allograft and autograft groups.

Data Synthesis

A qualitative synthesis of the evidence describes the comparison between allografts and autografts in PCL reconstruction. No statistical analysis could be performed, and no results were pooled in a meta-analysis due to clinical outcomes and methodologic heterogeneity.

Results

Study Selection and Characteristics

The PRISMA diagram (Fig 1) shows the study selection process. Four hundred ninty-two records were identified and screened resulting in 55 full-text articles assessed for eligibility. Fifty three articles were excluded because of wrong study design, wrong intervention, wrong comparison, wrong language, or wrong publication type. Two of the RCTs included compared allografts to autografts in PCL reconstruction. ^{25,26} Different types of allografts were used; 1 study reported the use of 4-stranded, y-irradiated tibialis anterior tendon allografts, ²⁵ and the other study used Achilles tendon and anterior tibial tendon allografts. ²⁶

Table 2 summarizes the data extraction and study characteristics of the 2 included studies, in which 145

patients from both studies were included; 30 patients received hybrid grafts, and 7 were lost to follow-up. The analysis included 108 patients: n=50 in the allograft group and n=58 in the autograft group. The mean follow-up time was 2.8 years in the study by Wang et al. and 5.6 years in the study by Li et al. The mean age in the groups was 32.2 (allografts) versus 31.3 (autografts) years and 30 (allografts) versus 29 years (autografts).

The study by Li et al.²⁵ reported inclusion criteria, such as persistent posterior laxity greater than or equal to grade 2, lack of response to medical and physical conservative treatment for more than 3 months, and PCL rupture identified by magnetic resonance imaging. The Wang et al. study²⁶ did not report any inclusion criteria. During the PCL reconstruction procedure, meniscal injuries were addressed, and patients received pre- and postsurgical rehabilitation such as physiotherapy. The studies reported patients' receiving the same postsurgical rehabilitation program; however, only the study by Li et al.²⁵ specified it as inclusive of a functional brace (6-12 weeks), passive range-of-motion exercises, progressive weight bearing after 2 weeks, and closed kinetic chain exercises. Only the study by Wang et al.²⁶ used radiographs for postoperative follow-up to measure the tibial and femoral tunnel widening, alignment of the knees and degenerative changes. The authors also reported a mild to moderate posterior laxity in 25% of patients for both groups, but the laxity grades were not quantified, nor was it reported how the laxity was assessed.

Quality Appraisal of the Evidence

Risk of bias assessment within PCL reconstruction studies

The two RCT studies were assessed for individual risk of bias according to the Cochrane Risk of Bias Tool.²⁴ Fig 2 represents the risk of bias as being unclear to high at the study level. The unclear risk was due to the incomplete description of the randomization process in the study by Li et al.²⁵ The risk for performance bias was unclear in both studies because of the lack of

 Table 2. Data Extraction Table

	Author, year	
	Li 2016 ²⁵	Wang 2004 ²⁶
Country	China	Taiwan
Interventions/Products	4-stranded y-irradiated tibialis anterior tendon allograft	Allograft (incl. Achilles tendon and anterior tibial tendons)
Comparator	Gracilis and semitendinosus tendon autograft Hybrid graft (y-irradiated tibialis anterior tendon allograft and semitendinosus tendon autograft)	Autograft (incl. quadriceps tendon- patellar bones and quadruple hamstrings)
Surgical procedure	Arthroscopic single-bundle PCLR	Arthroscopic single-bundle PCLR
Study design	RCT	RCT
Number of patients	90 [*]	55
	30 vs 30 vs 30	23 vs 32
Inclusion criteria	Persistent posterior laxity greater than or equal to grade 2, Lack of response to conservative treatment (medication and physical therapy) for more than 3 months Objective evidence of PCL rupture by MRI.	NR^{\dagger}
Exclusion criteria	Concomitant injury to other knee ligaments, previous surgery on the injured knee, or articular cartilage lesions greater than Outerbridge grade II based on the preoperative MRI scan and diagnostic arthroscopic examination	Patients with PCL avulsion fracture and combined ligament injuries were not included.
Rehabilitation (before or after PCLR)	The 3 groups of patients followed the same postoperative rehabilitation program. Physiotherapy Rehabilitation outside the hospital was not controlled.	Patients received the same postoperative rehabilitation (incl., e.g., exercise, functional knee braces, etc.). Physiotherapy: NR
Age of patients, mean \pm SD	$32.2 \pm 7.8 \; (20-40) \; \text{vs} \; 31.3 \pm 6.2 \; (20-40) \; \text{vs} \; 30.6 \pm 7.5 \; (20-40) \; \text{vs} $	30 \pm 12 (16-64) vs 29 \pm 12 (16-54)
(range), yrs	40) [§]	
Sex, n female (%) Further relevant patient characteristics	11 (40.7) vs 9 (34.6) vs 10 (37) ⁸ Patients with meniscal tears underwent partial meniscectomies (if irreparable) or repairs. No. of patients with no treatment of meniscal tears: 10/30 patients (33.3%) vs 9/30 patients (30%) vs 9/30 patients (30%); diff. n. s.	7 (30.4) vs 7 (21.9) Associated injuries were treated accordingly before PCLR. Concomitant treatments 5 meniscectomies 3 meniscal repairs 3 debridements for chondral lesions
Mean follow-up in yrs \pm SD	Overall [¶] : 5.6	Overall#: 2.83
	$5.7 \pm 0.3 \text{ vs } 5.5 \pm 0.2 \text{ vs } 5.6 \pm 0.5$	2.83 vs 2.75
Loss to follow-up, n (%)	Overall: 10 (11) 3 (10) vs 4 (13.33) vs 3 (10)	NR**
Patients included in analysis, n	27 vs 26 vs 27	23 vs 32
Effectiveness Outcomes Patient-reported function, activity levels and symptoms		
Lysholm score, mean \pm SD (range)	Preoperative: 64.1 ± 10.8 vs 63.8 ± 11.2 vs 62.3 ± 12.9 ; diff. n. s. with $P = 0.721$ Postoperative ^{††} : 85.2 ± 3.9 vs 87.8 ± 3.6 vs 86.9 ± 4.3 ; diff. n. s., with $P = 0.193$	Preoperative: NR Postoperative: ‡ 92.3 \pm 6.8 vs. 87.8 \pm 9.6; diff. n. s., with P = 0.077
Tegner score, mean \pm SD	Preoperative: $2.6 \pm 1.1 \text{ vs } 2.7 \pm 1.2 \text{ vs } 2.9 \pm 1.3; \text{ diff. n. s., with P} = 0.662$ Postoperative ^{††} $6.2 \pm 1.7 \text{ vs } 6.8 \pm 1.1 \text{ vs } 6.5 \pm 1.8; \text{ diff. n. s., with P} = 0.096$	4.70 ± 1.66 vs 4.73 ± 1.66 ; diff. n.s., with P = 0.976
IKDC score (subjective) , mean \pm SD (range)	Preoperative: $65.9 \pm 9.3 \text{ vs } 66.5 \pm 10.1 \text{ vs } 65.5 \pm 11.5; \text{ diff. n. s., with } P = 0.586$ Postoperative ^{††} $80.2 \pm 6.8 \text{ vs } 83.5 \pm 6.3 \text{ vs } 82.8 \pm 5.7; \text{ diff. n. s., with } P = 0.153$	NR

Table 2. Continued

	Author, year	
	Li 2016 ²⁵	Wang 2004 ²⁶
Clinical knee stability		
Lachman test (grade 0-1), n (%)	NR	Reverse Lachman Preoperative: NR Postoperative ^{‡‡} : mean \pm SD (range) ^{§§} : 0.70 ± 0.56 (0-2) vs 0.75 ± 0.67 (0-3); diff. n. s., with P = 0.898
Pivot shift test (grade 0-1), n (%)	Reverse Pivot shift (postoperative at final FU): 26 (96.3) vs 26 (100) vs 27 (100); diff. n. s., with $P = 0.407$	NR
Side-to-side difference in mm, mean \pm SD (range)	Measured with Instrumented Anteroposterior Measurements Postoperative: 3.5 \pm 1.1 (NR) vs 2.1 \pm 1.0 (NR ¶) vs 2.6 \pm 1.2 (NR #); diff. s. s., with P < 0.001	Measured with the KT-1000 arthrometer Preoperative: NR Postoperative***: $2.83 \pm 1.70 \ (1-6) \ vs \ 3.16 \pm 2.60 \ (1-10); \ diff. n. s., with P = 0.605$
IKDC score (objective; normal and nearly normal), n (%) ^{†††}	Postoperative (at Final FU): 24 (88.9) vs 25 (96.2) vs 25 (92.6); diff. n. s., with P = 0.716	Preoperative: NR Postoperative ^{†‡‡} : 14 (60.9) vs 23 (71.9); diff. n. s., with P = 0.391
Safety Outcomes		
Reoperation rate, n (%)	The study stated that no patient needed additional surgery because of recurrent or residual posterior laxity.	NR
Overall complications, n (%)	NR	0 (0) vs 7 (21.9)
Infection, n (%)	The study stated that no postoperative infection occurred.	0 (0) vs 2 $(6.3)^{\$\$}$
Other complications	Regarding complications, the study reported that there were no cases of major neurovascular, infectious, vascular, deep venous thrombosis, or wound complications.	Donor site pain: 0 (0) vs 4 (12.5) Reflex sympathetic dystrophy (RSD): 0 (0) vs 1 (3.1) Tibial and femoral tunnel Enlargement was comparable between the allograft and autograft groups with no significant difference at final follow-up: $ \textbf{Tibial}^{\ddagger\ddagger}: 12 \pm 20 \text{ (range: 0-90) vs } 12 \pm 14 \text{ (range: 0-43), with P} = 0.64 $ $ \textbf{Femoral}^{\ddagger\ddagger}: 5.3 \pm 22 \text{ (range: 0-50) vs } 13 \pm 19 \text{ (range: 0-55), with P} = 0.771 $

^{*116} points were assessed for eligibility. Of those, 90 points were randomized in the 3 treatment groups.

[†]The study reported only that it consisted of patients who underwent PCL reconstruction. In this context, the study states that the indication for surgery included pain and instability because of high-energy posterior cruciate ligament injury with failure of conservative treatments for 3 months.

[‡]Postoperative physiotherapy may have included, but was not limited to, the following: a functional brace (6-12 weeks), passive range of motion exercises, progressive weight bearing after 2 weeks, and closed kinetic chain exercises.

[§]The study reported on demographic data using the analyzed (not the enrolled) patients. Therefore, the denominator(s) are the analyzed patients within the groups.

Differences between treatment groups were not reported.

The overall mean was calculated by the review authors based on the mean FU time of the respective treatment groups.

[&]quot;The study only reported on the FU time in months. Overall: $34\pm10~(34\pm11~vs.~33\pm12)$

^{**}The study did not adequately report on the enrolment process. Therefore, the loss to follow-up rate could not have been calculated.

^{††}There were significant differences (p<0.05) between baseline scores and the last follow-up scores after surgery.

^{‡‡}The time point was not clearly reported.

The study reported only on the mean and range of the Lachman test.

^{||||95%}CI was reported: 3.083-3.9200.

^{¶95%}CI was reported: 1.6946-2.4941.

^{##95%}CI was reported: 2.1028-3.0397.

^{***}The time point was not clearly reported.

^{†††}The overall IKDC score has 4 qualitative characteristics regarding knee functionality: normal, nearly normal, abnormal, and severely abnormal. In this assessment, only the number of patients with normal and nearly normal characteristics were extracted and summed up.

^{‡‡‡}Femoral and tibia tunnel enlargements are presented in percentage of the tunnel width at follow-up over the width postoperative in AP view X-ray.

^{§\$\$}Excluding patients with hybrid grafts. The number of patients refers to the analysed patients, not the enrolled ones.

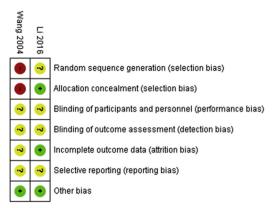


Fig 2. Risk of bias for studies comparing allografts to autografts in PCL reconstruction. This figure was created with the Review Manager, version 5.3 software to summarize the risk of bias at the study level.

reporting of the blinding of participants and the outcomes assessor, resulting in potential bias in the patient-reported outcomes. The risk of detection bias and reporting bias was unclear in both studies; no reporting protocol was published to indicate the complete range of outcomes. It was also unclear whether outcomes assessors were informed of the allocated intervention. The study by Wang et al. 26 was judged as being at a high risk of bias due to serious limitations in the randomization process. Graft randomization was done according to hospital admission, and the authors also reported issues with allograft availability, resulting in only 23 patients in the allograft group versus 32 in the autograft group. The lack of the allocation concealment resulted in a potential quasi-randomized study classification.

GRADE

The strength of the evidence in the 2 included studies was assessed according to GRADE²²; the evidence profile is provided in Table 3. The strength of the evidence for the effectiveness and safety of allografts compared to autografts in PCL reconstruction was judged as being low to very low. Evidence judged as low according to GRADE indicates that the authors' confidence in the effect estimate is questionable, and the true effect may be substantially different from the estimated effect. If evidence is judged as very low, evidence is either unavailable or does not permit a conclusion.²²

Results of Studies

Effectiveness outcome

The effectiveness of allografts compared to autografts in PCL reconstruction was measured by the crucial outcome of patient-reported function, activity levels and symptoms.

This outcome was measured in both studies by using the following instruments: The Lysholm score analyzed a total of 108 patients (excluding 27 patients with hybrid grafts), including 50 patients in the allografts group and 58 in the autografts group. The mean post-operative Lysholm score in the study by Li et al. was 85.2 \pm 3.9 in the allograft group (n = 27) and 87.8 \pm 3.6 in the autograft group (n = 26). In the study by Wang et al., the mean postoperative Lysholm score was 92.3 \pm 6.8 in the allograft group (n = 23) and 87.8 \pm 9.6 in the autograft group (n = 32). No statistically significant difference in the mean postoperative Lysholm score was identified between treatment groups (P > 0.05). The results were judged with very low certainty according to the GRADE methodology.

For the Tegner score, a total of 108 patients (excluding 27 patients with hybrid grafts) were analyzed into the allografts group (n = 50) or the autografts group (n = 58). The mean postoperative Tegner score in the study by Li et al. 25 was reported as 6.2 ± 1.7 in the allograft group (n = 27) and 6.8 ± 1.1 in the autograft group (n = 26). In the study by Wang et al., 26 it was reported as 4.7 ± 1.7 in the allograft group (n = 23) and 4.7 ± 1.7 in the autograft group (n = 32). No statistically significant difference in the mean postoperative Tegner score was identified between the groups (P > 0.05). The results were judged to be of low certainty according to the GRADE methodology.

The subjective IKDC score was reported only in the study by Li et al. ²⁵ A total of 53 patients (excluding 27 patients with hybrid grafts) were analyzed: the allografts group (n = 27) and the autografts group (n = 26) group. The mean postoperative subjective IKDC scores were 80.2 ± 6.8 and 83.5 ± 6.3 , respectively. The study did not report a statistically significant difference between the treatment groups (P > 0.05). This result was judged to be of low certainty according to the GRADE methodology.

Clinical knee stability was assessed using the reverse Lachman test, the reverse Pivot shift test and the instrumented side-to-side difference. Both included studies^{25,26} reported on the side-to-side difference in mm. A total of 108 patients (excluding 27 patients with hybrid grafts) were analyzed and placed into the allografts group (n = 50) or the autografts group (n = 58). The study by Li et al.²⁵ reported postoperative side-toside differences in mm of 3.5 \pm 1.1 and 2.1 \pm 1.0 in the allograft and autograft group, respectively. A statistically significant difference of 1.4 mm (P < 0.001) favoring the autografts group, with less side-to-side difference, was reported. This result was judged to be of very low certainty according to the GRADE methodology. The study by Wang et al., 26 however, reported no statistically significant difference between the groups: 2.8 ± 1.7 versus 3.2 ± 2.6 (P > 0.05) as measured by the KT-1000 arthrometer.

Table 3. GRADE Evidence Profile: Efficacy and Safety of Allografts in PCL Reconstruction

		Certainty Assessment A									
Number of Studies	Study Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Other Considerations	Allograft	Autoraft	Effect	Certainty	Importance
Effectiveness Patient-reported function, activity level and symptoms (follow-up: mean >2 years; assessed with: Lysholm score)											
Patient-reported function, activity level and symptoms (follow-up: mean > 2 years;	Randomized trials	Very serious**†	Serious [‡]	Not serious	Not serious	None	50	58	None of the studies found a statistically significant difference in the Lysholm score between treatment groups postoperatively. Li et al. 25 : 85.2 \pm 3.9 vs 87.8 \pm 3.6, diff. n. s. with P > 0.05 Wang et al. 26 : 92.3 \pm 6.8 vs 87.8 \pm 9.6, diff. n. s. with P > 0.05	⊕⊖⊖ Very low	Critical
assessed by Tegner score) 2 ^{1,2,25,26} Patient-reported function,	Randomized trials	Very serious* [†]	Not serious	Not serious	Not serious	None	50	58	None of the studies found a statistically significant difference in the Tegner score between treatment groups postoperatively. Li et al. 25 : 6.2 \pm 1.7 vs 6.8 \pm 1.1, diff. n. s. with P > 0.05 Wang et al. 26 : 4.7 \pm 1.66 vs. 4.73 \pm 1.66, diff. n. s. with P > 0.05	##©© Low	Critical
activity level and symptoms (follow-up: mean 5.6 years; assessed with: subjective IKDC score) 1 ^{2,25}	Randomized trials	Serious [†]	Not serious	Not serious	Serios [§]	None	27	26	Mean postoperative subjective IKDC score: 80.2 ± 6.8 vs 83.5 ± 6.3 , diff. n. s. with P > 0.05	⊕ ⊕ ○ ○ Low	Critical

		Certainty A	ssessment					ber of Patients*			
		Risk of				Other					
Number of Studies	Study Design	Bias	Inconsistency	Indirectness	Imprecision	Considerations	Allograft	Autoraft	Effect	Certainty	Importance
Clinical knee stability (follow- up: mean 2.8 years assessed by the Reverse Lachman test)											
1 ²⁶ Clinical knee stability (assessed	Randomized trials	Very serious*	Not serious	Not serious	Not serious	None	23	32	The study did not find a statistically significant difference in the reverse Lachman test postoperatively. Mean postoperative reverse Lachman test score: 0.7 ± 0.56 vs 0.75 ± 0.67 , diff. n. s. with $P > 0.05$	⊕⊕⊜ Low	Important
with: Reverse Pivot shift test)											
1 ²⁵ Clinical knee stability (follow-	Randomized trials	Serious [†]	Not serious	Not serious	Not serious	None	27	26	Postoperative reverse Pivot shift (Grade 0-1): $26/27$ (96.3%) vs $26/26$ (100%), diff. n. s. with $P > 0.05$	⊕⊕⊕⊜ Moderate	Important
up: mean >2 years; assessed with: KT arthrometer; better indicated by lower values)											
2 ^{25,26}	Randomized trials	Very serious* [†]	Serious	Not serious	Not serious	None	50	58	1 study found a statistically significant difference favoring autografts, while another study did not find any statistically significant difference based on the side-to-side difference measured by an instrumented knee laxity test. Side-to-side difference in mm: Li et al. 25 : 3.5 \pm 1.1 vs 2.1 \pm 1, diff. s. s. with P < 0.001 Wang et al. 26 : 2.83 \pm 1.7	⊕⊖⊖ Very low	Important

Table 3. Continued

		Certainty A	assessment					ber of l Patients*			
		Risk of				Other					
Number of Studies	Study Design	Bias	Inconsistency	Indirectness	Imprecision	Considerations	Allograft	Autoraft	Effect	Certainty	Importance
Clinical knee stability (follow- up: mean >2 years; assessed by objective IKDC)									vs 3.16 ± 2.6 , diff. n. s. with P > 0.05		
2 ^{25,26} Safety	Randomized trials	Very serious* ^{*/†}	Not serious	Not serious	Not serious	None	50	58	None of the studies found a statistically significant difference in the objective IKDC score between treatment groups postoperatively. Objective IKDC (normal and nearly normal): Li et al. ²⁵ : 24/27 (88.9%) vs 25/26 (96.2%) Wang et al. ²⁶ : 14/23 (60.9%) vs 23/32 (71.9%)	⊕⊕⊖⊖ Low	Important
Reoperations (follow-up: mean 1 ²⁵	1 5.6 years) Randomized trials	Not serious	Not serious	Serious ^{2,¶}	Serious#	None	27	26	The study stated that no patient needed additional surgery because of recurrent or residual posterior laxity: 0% vs 0% (further information: NR).	⊕⊕⊜⊝ Low	Critical
2 ^{25,26}	Randomized trials	Serious	Not serious	Not serious	Serious [§]	None	50	58	(reported in 1 study ²⁶): 0/23 (0) vs 7/32 (21.9%). Infections: Li et al ²⁵ : no postoperative infection; Wang et al. ²⁶ : 0/23 (0) vs 2/32 (6.3%; 1 acute and 1 late infection) Deep venous thrombosis (reported in 1 study ²⁵): 0/27 (0) vs 0/26 (0) Further reported complications: Donor site symptoms in	⊕⊕⊖⊖ Low	Critical

	Certainty Assessment							ber of Patients*			
		Risk of				Other					
Number of Studies	Study Design	Bias	Inconsistency	Indirectness	Imprecision	Considerations	Allograft	Autoraft	Effect	Certainty	Importance
									1 study 26 : 0/23 (0) vs 4/32 (12.5%); Reflex sympathetic dystrophy in 1 study 26 : 0/23 (0) vs 1/32 (3.1%) Li et al. 25 further stated narratively that no postoperative infection, no deep venous thrombosis, no cases of major neurovascular, infectious, vascular, deep venous thrombosis or wound complications in any of the 80 analyzed patients (of whom 27, 26 and 27 received allografts, autografts and hybrid grafts, respectively) occurred. Tibial and femoral tunnel enlargement in 1 study 26 : Tibial: 12 ± 20 (range: 0-90) vs 12 ± 14 (range: 0-43), n. s. with $P = 0.64$ Femoral: 5.3 ± 22 (range: 0-50) vs 13 ± 19 (range: 0-55), n. s. with $P = 0.771$		

CI, confidence interval; IKDC, International Knee Documentation Committee; MD, mean difference.

^{*}In Wang et al., there was high risk of bias for selection bias and unclear risk of bias as to whether patients were blinded.

[†]In Li et al., it was unclear whether the random sequence generation was adequate (selection bias), and whether patients were blinded.

[‡]None of the studies showed any statistically significant differences postoperatively in the Lysholm score between the allograft and the autograft groups. The non-statistically significant differences were not unanimously higher in one treatment group. This may be an indicator for heterogeneity. A calculation of the i-square is further needed to adequately assess how heterogeneous the results for this outcome may be.

[§]The optimal information size may have not been reached.

Heterogeneity may have been present because study results were not unanimous. Further calculation of the i-square is needed to elaborate on the extent of the heterogeneity.

The study referred to the patients who did not need additional surgery because of recurrent or residual posterior laxity. It was unclear to the review authors whether this refers to the overall reoperations rate or only the patients with recurrent or residual posterior laxity.

^{*}The optimal information size may have not been reached.

The objective IKDC score was reported in both studies. A total of 108 patients (excluding 27 patients with hybrid grafts) were analyzed in the allografts (n=50) or autografts (n=58) groups. No statistically significant difference (P>0.05 for both groups) in the post-operative objective IKDC score between treatment groups were reported. This result was judged to have low certainty according to the GRADE methodology.

Only the study by Wang et al. 26 reported on the reverse Lachman test. No statistically significant difference was reported in the mean postoperative reverse Lachman test scores (P > 0.05). This result was judged to have low certainty according to the GRADE methodology.

The reverse Pivot shift test was reported in the study by Li et al.²⁵ only. A total of 53 patients (excluding 27 patients with hybrid grafts) were analyzed in the allografts group (n = 27) and the autografts group (n = 26). No statistically significant difference was reported for the mean postoperative reverse Pivot shift test score (P = 0.4). This result was judged to have moderate certainty according to the GRADE methodology.

None of the studies reported on the health-related quality-of-life scores or the patient-satisfaction scores.

Safety outcome

The crucial outcomes to evaluate safety were graft failures, reruptures, reoperations and revisions, complications, and procedure-related mortality. The findings are summarized in Table 2. Reoperation was qualitatively reported only in the study by Li et al.²⁵ No quantitative data was reported, and the study stated only that no additional surgery was required as a result of recurrent or residual posterior laxity. This outcome was judged to have low certainty according to the GRADE methodology.

Graft failure, rerupture, revision rates, and procedure-related mortality were not reported in the included studies.

For the overall complications, only the study by Wang et al.²⁶ reported that no complications had occurred (0) in the allograft group versus 7 complications (21.9%) in the autograft group. This difference was not statistically tested, and the P value was not reported.

Both studies reported on the occurrence of infections. The study by Li et al.²⁵ reported no postoperative infections, whereas the study by Wang et al.²⁶ reported 2 infections (6.3%) in the autograft group. No infections occurred in the allograft group; the difference was not statistically tested, and no P value was reported. The study by Li et al.²⁵ reported that no deep-vein thrombosis occurred.

Arthrofibrosis, effusion, tenderness, hypoesthesia, and synovitis were not reported in either study.

The study by Wang et al.²⁶ reported the occurrence of donor-site symptoms: 0 in the allograft group and

12.5% in the autograft group. Reflex sympathetic dystrophy was also reported in 1/32 patients (3.1%) in the autograft group versus in 0/23 patients (0) in the allograft group. The differences were not statistically tested, and the the P value was not reported. The tibial and femoral tunnel widening was also reported, but no statistically significant difference between the allograft and autograft groups were found (P = 0.640; P = 0.771).

The study by Li et al.²⁵ reported no cases of major neurovascular, infectious, vascular, or wound complications and, overall, the outcome of complications was judged to have low certainty according to the GRADE methodology.

Discussion

The results of this study suggest that allografts and autografts may be comparable in PCL reconstruction regarding the crucial effectiveness outcome of patientreported function, activity level and symptoms. No statistically significant postoperative differences were reported between allografts and autografts for the Lysholm scale, the Tegner score, the subjective IKDC scores, or knee stability. The evidence base consisted of only 2 RCTs with unclear to high risk of bias and inconsistency and heterogeneity of reported outcomes and various graft sources. This is in line with results published in other systematic reviews^{14,27-29}; however, the certainty and strength of the evidence was judged according to the GRADE methodology only in this review. The overall strength of this evidence was judged to be low to very low, highlighting the limited confidence in reported effects estimates because the true effects might be substantially different from the estimated effects. Caution in interpretation of the results from the primary studies is, therefore, recommended.

The study by Li et al.25 reported a statistically significant difference of 1.4 mm (P < 0.001) favoring the autografts group; there was less side-to-side difference as measured by instrumented anteroposterior laxity. It was judged with a very low certainty of the evidence according to the GRADE. This, however, was not reported in the study by Wang et al., 26 which used the KT-1000 arthrometer to measure PCL laxity. The authors did, however, report a qualitative mild to moderate laxity in about 25% of patients in both the allografts and the autografts groups. The result should be interpreted with caution because there is potential inconsistency in the way the outcome was measured, and there exists the risk of reporting bias and heterogeneity of the graft types. The reported statistical mean difference was small, and the clinical relevance³⁰ is questioned due to the lack of a predefined, minimal clinically important difference (MCID) in PCL reconstruction. 14,15 A difference from the contralateral healthy side of less than 3 mm is considered normal; a

range from 3 to 5 mm is considered a gray zone; and a difference greater than 5 mm is considered to be a pathologic value regarding laxity due to anterior cruciate ligament ruptures. 31-33 It is, therefore, not plausible that a value smaller than 3 mm is a clinically relevant difference regarding the PCL, as outlined in the study by Li et al.²⁵ Patients might be less likely to report subjective instability or functional disability because the knee might be perceived as being stable as a result of the muscle support during functional activities. This might explain the reported success with conservative treatment for grade I and II PCL injuries due to the strong focus on physical rehabilitation. Autograft is well known for donor-site morbidity and concern for weakening the extensor and flexor mechanisms,² vet only the study by Wang et al.26 reported the measurement of isokinetic quadriceps and hamstring muscle strength and endurance. Postoperative concentric and eccentric muscle strength and endurance deficits were reported for both groups, but no statistically significant difference was found between the postoperative allograft and autograft scores.²⁶ preoperative values were not provided; comparisons were made to the uninjured side. The quantitative measurement of pre- or postoperative muscle strength was not reported or discussed as an outcome in other relevant studies. Perhaps adequate muscle strength and endurance have a more critical role to play and warrant further investigation and debate in the light of the increased risk of osteoarthritis development after PCL injury.1

Certain population subgroups such as high-level athletes⁹ might also be more vulnerable to muscle weaknesses following autografts and would benefit from specific, unbiased and well-designed studies.

One systematic review reported a higher level of postoperative activity in favor of autografts. Some reviews reported no significant differences in postoperative functional outcomes using autografts compared to allografts, whereas another meta-analysis reported insufficient evidence to report the superiority of allografts over autografts due to poor quality and data availability. 27,28

This review distinguishes the safety outcomes of PCL reconstruction from the effectiveness outcomes and suggests that there is insufficient evidence to judge superiority or inferiority of a particular graft type. It emphasizes poorly measured and reported complications because none of the primary studies reported on graft failure rates or revision or rerupture rates despite the reported graft failure rates of 11.6% and 12.6% mentioned in other studies. ^{5,6} Only the study by Wang et al. ²⁶ reported on an overall complication rate but without statistical testing. It seems that the reported complications are related to graft harvesting and do not capture all relevant complications. The study ²⁶ also

reported that allografts lead to fewer complications related to infections, donor-site pain and reflex sympathetic dystrophy in comparison to autografts. However, it was not reported how this was measured, nor was any statistical testing performed. It is also the only study to report that the femoral and tibial tunnels were comparable and that there was no statistically significant difference between the allografts and autografts. That is not mentioned in the other systematic reviews, despite its being indicated as a reason for graft failure.^{7,8} Y-irradiated allografts were included in the primary study by Li et al.²⁵ The authors acknowledged the limitations of this graft type, such as disease transmission, longer remodeling times and graft integration in the tunnels, as mentioned in other reviews.^{2,11,12} However, the main focus of this review was on measuring the stability outcomes and safety measures, such as no additional surgery needed due to posterior laxity, no cases of major infections and no vascular or wound complications, which were poorly measured and reported. No statistical testing was performed on group results, and a low GRADE judgment suggests caution in judging these findings because they might be substantially different from the true results.

The body of best evidence would benefit from future research focusing on studies using standardized, objective effectiveness- and safety-outcome measures, comparison with 1 type of graft only and methodologically sound, randomized study designs in PCL reconstruction.²⁸

Limitations

The primary limitation of this review is the inclusion of only 2 studies, which had a relatively small number of patients and a high risk of assessment bias (Fig 2). The studies reported a range of different allogenic graft types, such as y-irradiated, fresh-frozen and 4-stranded allografts and hamstring or quadriceps autografts. The extent of the potential effects of the differences and various disinfection procedures on clinical outcomes was unclear because each graft option presents unique characteristics that might affect the generalizablility of the comparison results. There is also poor reporting of complications, such as graft failures and reruptures resulting in variation, and there is heterogeneity in the reported outcomes, reflecting the lack of consensus on the criteria defining successful outcomes after PCL reconstruction. This resulted in the inability to pool results and, therefore, no meta-analysis could be performed. An MCID was not defined for the instrumented anteroposterior laxity measurement. Such a definition could be beneficial and would enable better judgment as to whether statistically significant differences found for objective, instrumented laxity outcomes are also clinically relevant³⁰ in practice.

Conclusions

Allografts may be comparable to autografts for crucial effectiveness outcomes, but insufficient evidence was found to judge crucial safety outcomes due to poor reporting of safety measures and outcomes. Results should be interpreted with caution because there is a lack of good-quality evidence to support the superiority of allografts over autografts; there is a high risk of bias in the primary studies and, overall, very low strength of the body of evidence according to the GRADE.

Acknowledgments

Project support by Information specialist: Tarquin Mittermayr, B.A. (Hons), M.A.

References

- 1. Sanders TL, Pareek A, Barrett IJ, et al. Incidence and longterm follow-up of isolated posterior cruciate ligament tears. *Knee Surg Sports Traumatol Arthrosc* 2017;25: 3017-3023.
- Johnson P, Mitchell S, Görtz S. Graft considerations in posterior cruciate ligament reconstruction. *Current Rev* Musculoskel Med 2018;11:521-527.
- MacDonald J, Rodenberg R. Posterior cruciate ligament injury. UpToDate, 2018, https://www.uptodate.com/ contents/posterior-cruciate-ligament-injury.
- **4.** Wang D, Graziano J, Williams RJI, Jones KJ. Nonoperative treatment of PCL injuries: Goals of rehabilitation and the natural history of conservative care. *Curr Rev Musculoskel Med* 2018;11:290-297.
- Hammoud S, Reinhardt KR, RG M. Outcomes of posterior cruciate ligament treatment: A review of the evidence. Sports Med Arthrosc Rev 2010;18:280-291.
- 6. Tucker CJ, Cotter EJ, Waterman BR, et al. Functional outcomes after isolated and combined posterior cruciate ligament reconstruction in a military population. *Orthop J Sports Med* 2019;7:2325967119875139-2325967119875139.
- Noyes FR, Barber-Westin S. Posterior ligament revision reconstruction, part1: Causes of surgical failure in 52 consecutive operations. Am J Sports Med 2005;33:646-654.
- 8. Zawodny S, Miller M. Complications of posterior cruciate ligament surgery. *Sports Med Arthrosc Rev* 2010;18: 269-274.
- 9. Devitt BM, Dissanayake R, Clair J, et al. Isolated posterior cruciate reconstruction results in improved functional outcome but low rates of return to preinjury level of sport: A systematic review and meta-analysis. *Orthopead J Sports Med* 2018;6:2325967118804478.
- **10.** Hoher J, Scheffler S. Graft choice and graft fixation in PCL reconstruction. *Knee Surg Sports Traumatol Arthrosc* 2003;11:297-306.
- 11. Friedberg R. *Anterior cruciate ligament injury*. UpToDate, 2018.
- **12.** Hudgens JL, Gilette BP, Krych AJ, et al. Allograft versus autograft in posterior cruciate ligament reconstruction: An evidence-based systematic review. *J Knee Surg* 2013;26: 109-115.

- Fahey M, Indelicto PA. Bone tunnel enlargement after anterior cruciate ligament replacement. Am J Sports Med 1994;22:410-414.
- **14.** Ansari AS, Dennis BB, Horner NS, et al. Influence of graft source on postoperative activity and joint laxity in posterior cruciate ligament reconstruction: A systematic review. *Arthroscopy* 2018:262-274.e266.
- Belk JW, Kraeutler MJ, Purcell JM, McCarty EC. Autograft versus allograft for posterior cruciate ligament reconstruction: An updated systematic review and metaanalysis. *Am J Sports Med* 2018;46:1752-1757.
- 16. Zeng C, Gao SG, Li H, et al. Autograft versus allograft in anterior cruciate ligament reconstruction: A meta-analysis of randomized controlled trials and systematic review of overlapping systematic reviews. *Arthroscopy* 2016;32: 153-163.e118.
- 17. European Network of Health Technology Assessment (EUnetHTA) Joint Action 2 WP. HTA Core Model version 3.0 (pdf). 2016.
- **18.** Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *J Clin Epidemiol* 2009;62:e1-e34.
- 19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int J Surg* 2010;8:336-341.
- 20. Goetz G, De Villiers C. Allograft for anterior and posterior cruciate ligament reconstruction: Systematic review. Decision support document 116. LBI-HTA 2019.
- European Network of Health Technology Assessment Joint Action 2 WP. Levels of evidence: Internal validity of randomised controlled trials 2015.
- **22.** Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64: 383-394.
- 23. Hohmann E, Feldman M, Hunt TJ, et al. Research pearls: How do we establish the level of evidence? *Arthroscopy* 2018;34:3271-3277.
- **24.** Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011:343:d5928.
- **25.** Li J, Kong F, Gao X, Shen Y, Gao S. Prospective randomized comparison of knee stability and proprioception for posterior cruciate ligament reconstruction with autograft, hybrid graft, and gamma-irradiated allograft. *Arthroscopy* 2016;32:2548-2555.
- **26.** Wang CJ, Chan YS, Weng LH, et al. Comparison of autogenous and allogenous posterior cruciate ligament reconstructions of the knee. *Injury* 2004;35: 1279-1285.
- 27. Tian P, Hu W, Li Z, et al. Comparison of autograft and allograft tendons in posterior cruciate ligament reconstruction: A meta-analysis. *Medicine (Baltimore)* 2017;96: e7434.
- **28.** Martin RK, Levy BA. Editorial commentary: What is the best graft source for postreior cruciate ligament reconstruction? *Arthroscopy* 2018;35:275-276.
- **29.** Lee YS, Lee SH, Lee OS. Graft sources do not affect to the outcome of transtibial posterior cruciate ligament

- reconstruction: A systematic review. *Arch Orthop Trauma Surg* 2018;138:1103-1116.
- **30.** Dhawan A, Brand JC, Provencher MT, et al. Research pearls: The significance of statistics and perils of pooling. *Arthroscopy* 2017;33:1099-1101.
- **31.** Huber FE, Harner C, Lephart S. Intratester and intertester reliability of the KT-1000 arthrometer in the assessment of posterior laxity of the knee. *Am J Sports Med* 1997;25:479-485.
- **32.** Daniel DM, Stone ML, Sachs R. Instrumented measurement of anterior knee laxity in patients with acute
- anterior cruciate ligament disruption. *Am J Sports Med* 1985;13:401-407.
- **33.** Anderson A, Snyder R, Federspiel C, Lipscomb A. Instrumented evaluation of knee laxity: A comparison of five arthrometers. *Am J Sports Med* 1992;20:135-140.
- 34. Liang Chen YW, Yufeng Wu, Dawei Gao, Guoping Xie. Autografts promote post-operative activity level compared to allografts in posterior cruciate ligament reconstruction: A systematic review and meta-analysis. *Int J Clin Exper Med* 2017;10:9807-9817.