NEJM Evidence

Published July 18, 2024
NEJM Evid 2024; 3 (8)
DOI: 10.1056/EVIDoa2400056

ORIGINAL ARTICLE

Operative versus Nonoperative Treatment of Proximal Hamstring Avulsions

Elsa Pihl, M.D., Ph.D.,[1,2] Sofia Laszlo, M.D.,[3] Anne-Mari Rosenlund, M.D.,[4,5] Målfrid Holen Kristoffersen, M.D., Ph.D.,[6] Jörg Schilcher, M.D., Ph.D.,[7,8] Carl Johan Hedbeck, M.D., Ph.D.,[1,2] Mikael Skorpil, M.D., Ph.D.,[9] Chiara Micoli, M.S.,[10] Martin Eklund, Ph.D.,[10] Olof Sköldenberg, M.D., Ph.D.,[1,2] Frede Frihagen, M.D., Ph.D.,[4,11] and Kenneth B. Jonsson, M.D., Ph.D.,[3] for the PHACT group*

Abstract

BACKGROUND Operative treatment is widely used for acute proximal hamstring avulsions, but its effectiveness compared with that of nonoperative treatment has not been shown in randomized trials.

METHODS In this noninferiority trial at 10 centers in Sweden and Norway, we enrolled patients 30 to 70 years of age with a proximal hamstring avulsion in a randomized trial and a parallel observational cohort. Treatments were operative reinsertion of the tendons or nonoperative management. The primary end point was the Perth Hamstring Assessment Tool (PHAT) at 2 years of follow-up. Secondary outcomes included scores on the Lower Extremity Functional Scale (LEFS).

RESULTS A total of 119 patients were enrolled in the randomized trial and 97 patients in the observational cohort. In the per-protocol analysis of the randomized trial, the mean (±standard deviation) PHAT scores were 79.9±19.5 and 78.5±19.4 in the operative and nonoperative groups, respectively (PHAT scores range from 0 to 100, with higher scores indicating higher function). The prespecified noninferiority limit of 10 points was not crossed (mean difference, –1.2; 95% confidence interval [CI], –8.6 to 6.2; P=0.009 for noninferiority). Analyses of secondary outcomes, including a mean difference in the LEFS score of –1.6 (95% CI, –5.2 to 2.0), aligned with the primary outcome. The observed numbers of adverse events in the randomized trial were nine in the operative group versus three in the nonoperative group (odds ratio, 0.3; 95% CI, 0.1 to 1.2). In the analysis of the observational cohort, the mean PHAT score difference between the nonoperative and operative treatment groups was –2.6 (95% CI, –9.9 to 4.6).

CONCLUSIONS In patients 30 to 70 years of age with proximal hamstring avulsions, nonoperative treatment was noninferior to operative treatment. (Funded by Afa Försäkring and others; ClinicalTrials.gov number, NCT03311997.)

A complete list of investigators in the PHACT group is provided in the Supplementary Appendix, available at evidence.nejm.org.

The author affiliations are listed at the end of the article.

Dr. Jonsson can be contacted at kenneth.jonsson@uu.se or at the Department of Surgical Sciences, Uppsala University Hospital, 751 85 Uppsala, Sweden.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

Introduction

Despite common beliefs in the superiority of surgery over conservative management, many operative procedures lack evidence of efficacy from randomized trials. This is particularly evident for numerous orthopedic procedures.[1,2] One such example is operative treatment for proximal hamstring avulsions (PHAs), the most severe hamstring injury. PHA often affects middle-aged, physically active individuals.[3,4] The injury is acutely painful; the most common injury mechanism is a slip and fall during sports or daily activities, and it often involves an involuntary front split. At long-term follow-up, many patients report persistent discomfort in their buttock, numbness, and stiffness in their affected leg as well as failure to return to their preinjury activity level.[5,6]

Operative treatment has been the recommended approach for PHA, supported predominantly by retrospective case series and cohort studies, which reported favorable outcomes after surgery.[7,8] Consequently, operative treatment for PHA has increased,[4,9] and a previous meta-analysis, which included only a minor fraction of patients not operatively treated within the cohort, concluded that surgical repair was superior compared with nonoperative treatment.[7] Two recently published cohort studies and one case-control study compared the outcomes of nonoperative and operative treatments of PHA. None found evidence in favor of operative treatment.[10-12] However, they all suffer from inherent methodologic limitations.

We conducted a randomized controlled trial of nonoperative treatment or surgical repair among adults 30 to 70 years of age with an acute PHA (PHACT [Proximal Hamstring Avulsion Clinical Trial]). Considering the cost and risks associated with surgery and the fact that most of the literature recommends operative treatment, the trial was designed to test for the noninferiority of nonoperative treatment.

Methods

TRIAL DESIGN

PHACT was a multicenter, preference-tolerant, randomized controlled noninferiority trial conducted at eight university and regional hospitals in Sweden and two university hospitals in Norway. The protocol has been published previously[13] and is provided, along with the statistical analysis plan, with the full text of this article at evidence.nejm.org. Ethical approval was granted by the Uppsala Regional Ethical Committee and by the Regional Committee of Medical and Health Research Ethics in Norway. The trial was designed by the authors, with data collection undertaken by consortium members. The authors are responsible for accuracy of the data, the trial's adherence to the protocol, manuscript writing, and the decision to publish. AFA Försäkring (an organization owned by Sweden's labor market parties) was the main sponsor of the trial, but it had no role in protocol development, data analysis, or manuscript preparation. Details about the participating sites, eligibility criteria, treatment allocation, interventions, follow-up, outcome methodology, and statistical analysis are provided in the Supplementary Appendix.

PARTICIPANTS

Patients 30 to 70 years of age with a suspected PHA were assessed for eligibility at the trial sites. Eligibility required at least a moderately active lifestyle and magnetic resonance imaging (MRI) that confirmed an acute (within 4 weeks) and complete avulsion of at least two of the three tendons. A moderate activity level was defined as at least regular participation in strenuous work or leisure activities (e.g., vigorous walking, running, skiing, ladder climbing, heavy gardening). Patients with multiple injuries or comorbidities such as severe obesity (body-mass index [BMI] >35 kg/m^2) or diabetes with secondary complications were excluded due to increased perioperative and postoperative risks. Additional criteria and the rationale behind them are described in the Supplementary Appendix, Inclusion and Exclusion. The demographic characteristics of the trial population are further outlined in Table S10.

TRIAL PROCEDURE

Eligible patients were informed about the trial and asked to participate in the randomized controlled trial. Those who enrolled were randomly assigned in a 1:1 ratio to either operative reattachment of the tendons or to nonoperative treatment. In cases in which the patient and the surgeon could not reach equipoise to treatment options as part of their shared decision-making process, the option to enroll in a parallel observational cohort was offered. This cohort provided identical treatment options and adhered to the same follow-up protocol.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

For operative treatment, perioperative and postoperative antibiotic and anticoagulant treatments were guided by local policy. The choices of skin incision, dissection, closure, and wound dressing were left to the discretion of the surgeon. The tendons were reattached using at least two suture anchors. The physiotherapy protocol was identical to that of the nonoperatively treated patients who initiated their rehabilitation immediately after trial inclusion.

Participants, surgeons, and research coordinators who collected patient outcome assessments were aware of the assigned treatment group. Patients were asked to wear clothing that covered surgical scars to facilitate blinding of the physiotherapist performing the functional tests.

OUTCOMES

The primary outcome was the Perth Hamstring Assessment Tool (PHAT) score at 24 months' follow-up. PHAT is a condition-specific questionnaire with scores ranging from 0 to 100, with a higher score corresponding to higher function. The score is the only outcome measure validated for follow-up after PHA,[5,14-17] and it was, together with secondary patient-reported outcomes, obtained at baseline (using preinjury recall) and at 3, 6, 12, and 24 months. The preinjury recall measurements were obtained to estimate the degree of and time to full recovery.

Secondary patient-reported outcome measures included the Lower Extremity Functional Scale (LEFS), which is a tool for assessing functional status in populations with lower extremity musculoskeletal conditions (the maximum [best] score on LEFS is 80, and the minimum [worst] score is 0).[18] Information regarding physical activity was collected by using the short form of the International Physical Activity Questionnaire (IPAQ; higher scores indicate higher activity, with the lowest possible score being 0 and the highest possible score being 321).[19] In addition, self-reported data on general satisfaction, estimation of recovery, and return to sports were collected.

The physical performance was assessed at 6-, 12-, and 24-month follow-up visits through functional tests and by an isometric strength test using a dynamometer. Functional tests comprised the timed single-step test,[20] the single-leg hop, and the triple hop for distance tests.[21,22] All MRIs performed at each site for initial diagnosis were centrally re-examined by the trial radiologist (M.S.) to verify tendon avulsion and measure retraction (Fig. S1). MRI was also performed at the 24-month follow-up to evaluate hamstring muscle volume and fat infiltration in both thighs (Fig. S2). For MRI and functional test data, we calculated the limb symmetry index (LSI) expressed in percentages using the formula LSI percent equals (value of injured side/value of uninjured side) $\times 100$.

STATISTICAL METHODS

The sample size calculation was updated in a statistical analysis plan amendment in December 2018. Initially, the goal was to recruit 60 patients in the randomized trial to achieve 80% power for showing noninferiority using a margin corresponding to the minimal detectable change in the PHAT score (16 points). The inclusion rate surpassed expectations, with more than 50 patients enrolled within 14 months. We concluded that the injury's incidence had been underestimated and that trial logistics were effective. Thus, we decided to increase the sample size to be able to provide more precise answers to our research question. We decreased the noninferiority margin from 16 to 10, which represents about one half the expected standard deviation (SD) of the PHAT score and is smaller than the reported minimal detectable within-patient change (16 points) of the PHAT score[14,16] (details are provided in the Supplementary Appendix, Motivation of Extending the Sample Size and Increasing Power). We estimated that 60 patients in each randomized group were needed to achieve 74 to 90% power (depending on the assumed SD of the PHAT score) for showing noninferiority when assuming a dropout rate of 10%, an SD of the PHAT score in the range of 16 to 20, and a 2.5% one-sided alpha.

The per-protocol population (participants who were randomized to receive a specific treatment and who did receive it according to the protocol) was used for the primary analysis. PHAT scores were compared between the two treatment groups using a linear regression model, including an indicator of randomized treatment group assignment adjusted for age, sex, and degree of tendon retraction, and using inverse probability weighting to adjust for adherence to the assigned treatment. The probability weights were calculated based on age, sex, degree of tendon retraction, and treatment group. If the lower boundary for the two-sided 95% confidence interval (CI) in the absolute difference of the PHAT score in the nonoperative group and the operative group was greater than –10 points, the nonoperative treatment would be deemed to be noninferior. Missing data on individual PHAT questions were imputed using predictive mean matching with a donor pool of five. A summary

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

of the frequency of internally missing data points for the PHAT and LEFS scores is shown in Table S8.

The following prespecified sensitivity analyses were performed for the primary end point: analysis after model-based multiple imputation by chained equations of the PHAT scores for participants who missed follow-up appointments and analyses on the intention-to-treat population (participants analyzed in groups defined according to the treatment that they were randomized to) and the as-treated population (patients analyzed in groups defined according to the treatment that the patients actually received). In addition, because the PHAT score is constrained between 0 and 100, a prespecified sensitivity analysis was performed of the primary end point using beta regression.

Secondary outcomes were contrasted between the two treatment groups using regression models adjusting for the same variables as in the primary end point analysis. Logistic regression was used for binary end points (adverse events and return to sports), and linear regression was used for all other secondary end points. Because odds ratios are estimated from logistic regression models and because these odds ratios may overestimate relative risks, we performed a post hoc analysis estimating marginal relative risks for the adverse events and the return-to-sports secondary end points.

Prespecified subgroup analyses were performed according to age (<50 years of age vs. ≥50 years of age), sex, and degree of tendon retraction (<2 vs. ≥2 cm). Heterogeneity of treatment effects was assessed using an interaction term between treatment allocation and age, sex, and degree of tendon retraction, respectively, in the regression models.[23] We analyzed the observational cohort using propensity score adjustment by inverse probability weighting. The propensity score was based on age, sex, baseline IPAQ, and the degree of tendon retraction. P values are reported only for the primary outcome in the randomized trial. For secondary outcomes, two-sided 95% CIs are reported. These CIs were not adjusted for multiplicity and should not be used to judge clinical effects.

Results

TRIAL POPULATION

Between October 2017 and July 2020, a total of 216 patients were enrolled: 119 in the randomized trial and 97 in the observational cohort. The final follow-up was completed in September 2022. In the randomized trial, 58 patients were assigned to operative treatment, and 61 patients were assigned to nonoperative treatment; both treatment groups had two early crossovers. Among the 97 patients in the observational cohort, 44 (45%) were treated operatively, and 53 (55%) were treated nonoperatively (Fig. 1). Four of the 10 participating sites enrolled 71% of the participants in the randomized trial (Table S1). At the 24-month follow-up, 115 (97%) and 85 (88%) participants remained in the randomized and observational cohorts, respectively.

The baseline characteristics of patients in the randomized and observational cohorts were similar (Table 1). In the combined randomized and observational cohorts, 60% of patients were female, and 40% were male. Most patients sustained their injury during sporting activities (n=117 [54%]) and had all tendons avulsed greater than or equal to 2 cm (n=190 [90%]). The mean (±SD) tendon retraction was 5.0 ± 2.6 cm, and upon central reevaluation, all MRIs revealed avulsions compatible with the inclusion criteria.

PRIMARY OUTCOME

Primary outcome data (PHAT score at 24 months) were available for 114 (96%) of 119 patients in the randomized trial. In the primary per-protocol analysis, the between-group difference in the nonoperative treated group versus operative treated group was −1.2 (95% CI, −8.6 to 6.2; P=0.009 for noninferiority) (Table 2). This result is consistent with noninferiority as the lower bounds of the 95% two-sided CI for the differences between means in PHAT did not include the predefined noninferiority margin of −10 (Fig. 2). Our results were consistent across sensitivity analyses and the propensity score–adjusted analysis of the observational cohort (mean difference, −2.6; 95% CI, −9.9 to 4.6). Furthermore, evaluation for heterogeneity of treatment effects was not consistent with notable effect modification (Table S3).

SECONDARY OUTCOMES

In the per-protocol analysis of the randomized trial, the mean (±SD) PHAT score decreased from baseline to the 3-month visit: from 92.1 ± 11.2 to 60.9 ± 20.4 in the operatively treated group and from 91.7 ± 11.5 to 66 ± 20.3 in the nonoperatively treated group. The mean (±SD) PHAT score gradually increased, and at 24 months, it was 79.9 ± 19.5 in the operatively treated patients and

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

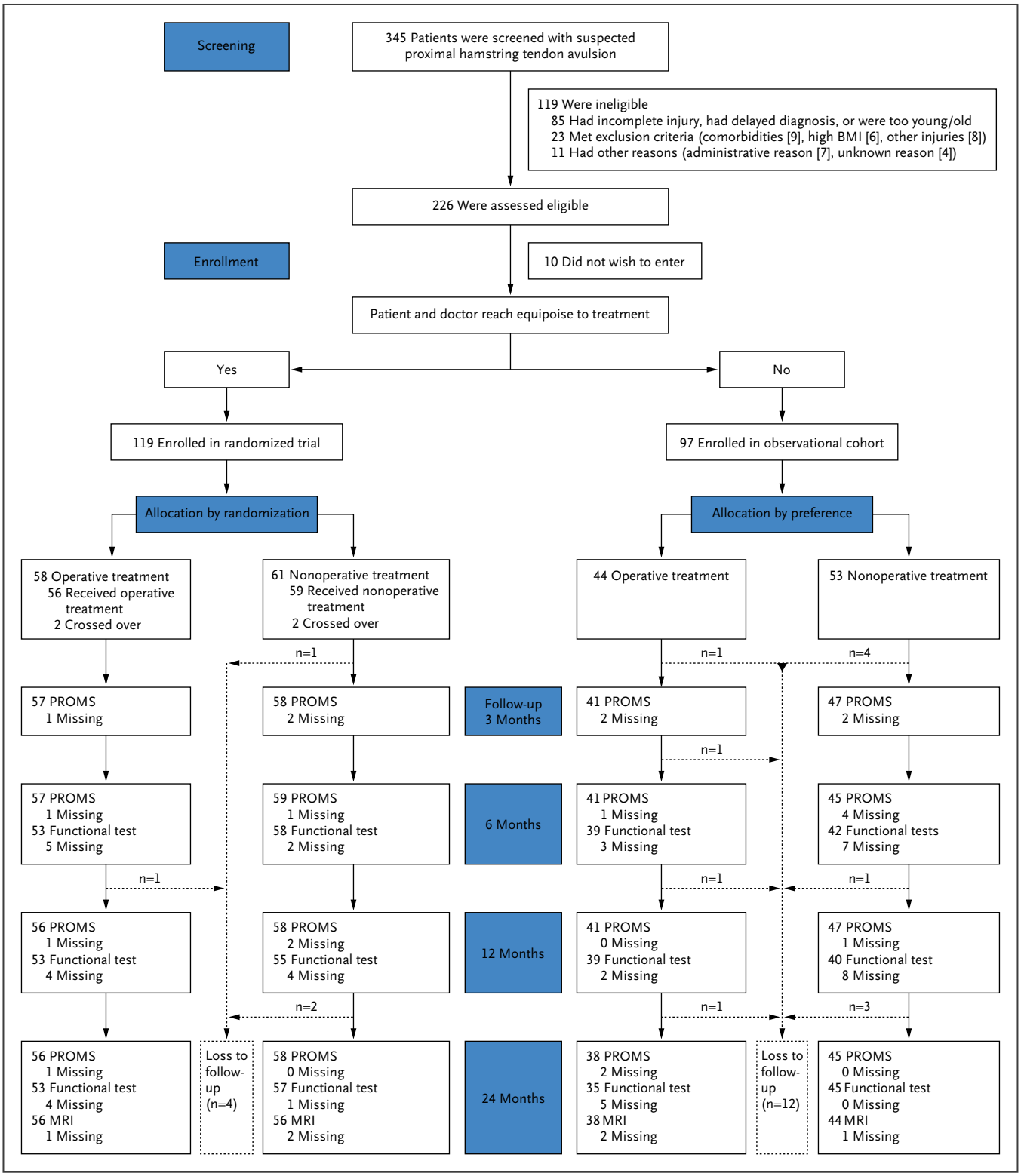


Figure 1. Screening, Randomization, and Follow-up of Participants.

Patient-reported outcome measurements (PROMS) included the Perth Hamstring Assessment Tool. Patients who were referred to one of the trial sites with a suspected proximal hamstrings avulsion were screened. Inclusion criteria included age, time since injury, and magnetic resonance imaging (MRI) verification of a complete avulsion of at least two tendons. The most common reasons for ineligibility were delayed diagnosis and incomplete tendon avulsions. Patients were enrolled in either the randomized trial or a parallel observational cohort. In the randomized trial, two patients from each treatment group crossed over. BMI denotes body-mass index.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

Table 1. Baseline Characteristics.*

	Randomized Trial (N=119)		Observational Cohort (N=97)	
Characteristics†	**Operative (n=58)**	**Nonoperative (n=61)**	**Operative (n=44)**	**Nonoperative (n=53)**
Age (yr)	54.4 (49.7–59.6)	53.4 (49.5–57.4)	49.8 (46.3–58.1)	55.1 (50.2–58.8)
Female sex	32 (55.2)	40 (65.6)	20 (45.5)	37 (69.8)
Injured side				
Right	25 (43.1)	34 (55.7)	17 (38.6)	32 (60.4)
Missing	0 (0.0)	0 (0.0)	2 (4.5)	2 (3.8)
Tendons avulsed‡				
Conjoint tendon only	1 (1.7)	1 (1.6)	0 (0.0)	2 (3.8)
All tendons	56 (96.6)	60 (98.4)	41 (93.2)	48 (90.6)
Missing	1 (1.7)	0 (0.0)	3 (6.8)	3 (5.7)
Tendon retraction (cm)‡	5.0 (3.0–7.50)	4.5 (3.5–6.5)	6.5 (4.0–7.5)	3.5 (2.0–6.4)
Missing	1 (1.7)	0 (0.0)	3 (6.8)	3 (5.7)
Time from injury to inclusion (days)	13.5 (10.0–17.7)	13.0 (9.0–18.0)	12.0 (9.0–18.2)	16.0 (9.0–20.0)
Time to surgery (days)	18 (13–22)		19 (13–24)	
Activity at injury				
Sporting activity	34 (58.6)	30 (49.2)	27 (61.4)	26 (49.1)
Winter skiing	12 (20.7)	7 (11.5)	6 (13.6)	7 (13.2)
Winter skating	5 (8.6)	4 (6.6)	1 (2.3)	2 (3.8)
Water skiing	3 (5.2)	5 (8.2)	6 (13.6)	4 (7.5)
Other sports	14 (24.1)	14 (23.0)	14 (31.8)	13 (24.5)
Slip or fall	16 (27.6)	26 (42.6)	8 (18.2)	18 (34.0)
Vehicle accident	1 (1.7)	3 (4.9)	3 (6.8)	2 (3.8)
Other	7 (12.1)	2 (3.3)	6 (13.6)	7 (13.2)
Employment status				
Employed	47 (81.0)	47 (77.0)	33 (75.0)	37 (69.8)
Self-employed	6 (10.3)	8 (13.1)	7 (15.9)	8 (15.1)
Unemployed	0 (0.0)	2 (3.3)	0 (0.0)	1 (1.9)
On sick leave	2 (3.4)	1 (1.6)	0 (0.0)	3 (5.7)
Retired	3 (5.2)	3 (4.9)	4 (9.1)	3 (5.7)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.9)
PHAT preinjury	97.0 (87.5–100.0)	95.0 (90.3–100.0)	100.0 (97.0–100.0)	93.5 (80.0–100.0)
Missing	3 (5.2)	3 (4.9)	1 (2.3)	3 (5.7)
LEFS preinjury	80.0 (78.0–80.0)	80.0 (78.0–80.0)	80.0 (78.7–80.0)	80.0 (77.0–80.0)
Missing	2 (3.4)	2 (3.3)	0 (0.0)	1 (1.9)
IPAQ preinjury	47.5 (33.3–93.5)	55.4 (28.1–95.9)	81.5 (47.7–113.2)	57.9 (38.2–97.3)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.8)
BMI	26.0 (23.8–29.1)	25.7 (23.2–28.6)	25.2 (23.4–28.1)	26.5 (23.9–28.3)
Missing	0 (0.0)	2 (3.3)	1 (2.3)	2 (3.8)
Diabetes	2 (3.4)	1 (1.6)	0 (0.0)	3 (5.7)
Cortisol treatment	4 (6.9)	1 (1.6)	1 (2.3)	3 (5.7)

* Values are presented as median (interquartile range) or n (%). The table contains no imputed data. BMI denotes body-mass index; IPAQ, International Physical Activity Questionnaire; LEFS, Lower Extremity Functional Scale; and PHAT, Perth Hamstring Assessment Tool.

† PHAT measures patient-reported pain and function during various activities, and scores range from 0 (worst) to 100 (best).[16] There was one patient completely missing the PHAT score. The other missing PHAT scores are due to single omitted items within the score. The maximum best score of LEFS is 80, and the minimum worst score is 0.[18] The short form of the IPAQ was used.[19] Metabolic equivalent of task (MET) hours were calculated as recommended (https://sites.google.com/view/ipaq/score). Higher MET hours indicate higher activity, with the lowest possible score being 0 and the highest possible score being 321.

‡ The magnetic resonance images performed for diagnosis were re-examined by a consultant in radiology to verify tendon avulsion and retraction. The degree of tendon retraction was measured by using the "direct retraction measurement" method according to van der Made et al.[24] Missing data are attributed to the inability to retrieve the original images for secondary evaluation.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

Table 2. Primary and Secondary Outcomes at 24 Months.*

Outcome†	Operative Treatment		Nonoperative Treatment		Between-Group Difference or Odds Ratio‡	Significance
	No. of Patients (Missing)	Mean±SD	No. of Patients (Missing)	Mean±SD	Difference (95% CI)	P Value (Noninferiority)
Primary outcome: total PHAT score						
PP	54 (2)	79.9±19.5	57 (2)	78.5±19.4	–1.2 (–8.6 to 6.2)	0.009
PP with imputation§	56 (0)	78.5±20.8	59 (0)	78.5±19.3	0.4 (–7.2 to 7.9)	—
ITT	56 (2)	80.4±19.3	58 (3)	77.7±20.0	–2.1 (–9.3 to 5.1)	—
ITT with imputation§	58 (0)	79.2±20.4	61 (0)	78.0±19.8	–0.5 (–7.8 to 6.7)	—
AS	55 (3)	79.1±20.2	59 (2)	79.0±19.2	–0.01 (–8.1 to 8.1)	—
OBS COHORT	38 (6)	81.1±17.3	45 (8)	82±16.1	–2.6 (–9.9 to 4.6)	—
Secondary outcomes: patient reported						
Total LEFS score						
PP	54 (2)	73.4±8.4	57 (2)	72.0±10.7	–1.6 (–5.2 to 2.0)	—
OBS COHORT	38 (6)	73.7±9.3	45 (8)	72.6±8.7	–2.2 (–5.9 to 1.5)	—
IPAQ short (MET hours)						
PP	54 (2)	61.7±54.9	57 (2)	63.2±52.1	3.0 (–19.0 to 25.0)	—
OBS COHORT	38 (6)	61.0±37.0	45 (8)	79.1±45.1	19.6 (–0.2 to 39.4)	—
Patient satisfaction: VAS (0–100)¶						
PP	54 (2)	83.2±23.1	57 (2)	76.7±23.5	–6.8 (–15.7 to 2.1)	—
OBS COHORT	38 (6)	85.1±15.3	44 (9)	80.5±25.6	–8.3 (–20.0 to 3.3)	—
Patient recovery: VAS (0–100)‖						
PP	54 (2)	80.8±18.9	57 (2)	74.9±23.7	–6.1 (–14.5 to 2.3)	—
OBS COHORT	38 (6)	81.1±18.4	45 (8)	80.5±18.9	–7.5 (–17.9 to 2.9)	—
Secondary outcomes: functional tests**						
Maximum isometric force (LSI%)						
PP	51 (5)	91.7±20.2	56 (3)	88.5±20.8	–4.5 (–12.3 to 3.4)	—
OBS COHORT	35 (9)	92.8±11.5	45 (8)	84.7±15.8	–10.9 (–18.5 to –3.4)	—
Single-step tests (LSI%)						
PP	50 (6)	100.4±6.9	55 (4)	101.5±13.0	1.0 (–3.2 to 5.2)	—
OBS COHORT	35 (9)	97.7±8.9	45 (8)	99.3±14.3	2.2 (–2.9 to 7.3)	—
Single-leg hop test (LSI%)						
PP	46 (10)	100.4±18.6	49 (10)	95.0±26.9	–6.6 (–16.1 to 2.8)	—
OBS COHORT	34 (10)	103.1±21.2	40 (13)	94.2±19.5	–11.7 (–21.7 to –1.6)	—
Triple-leg hop test (LSI%)						
PP	45 (11)	99.5±12.6	49 (10)	95.7±19.3	–4.6 (–11.2 to 2.0)	—
OBS COHORT	34 (10)	102.1±18.8	40 (13)	94.6±15.4	–9.5 (–18.5 to –0.4)	—
Secondary outcomes: MRI						
Hamstring muscle volume (LSI%)††						
PP	53 (3)	88.0±10.1	53 (6)	84.4±12.6	–4.6 (–8.6 to –0.6)	—
OBS COHORT	37 (7)	89.9±10.0	44 (9)	84.8±9.9	–6.5 (–10.8 to –2.1)	—
Fat fraction (LSI%)‡‡						
PP	53 (3)	136.6±30.0	53 (6)	152.4±40.4	20.5 (7.9 to 33.1)	—
OBS COHORT	37 (7)	128.7±33.4	44 (9)	145.4±39.5	29.9 (11.3 to 48.5)	—

(continued)

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

Table 2. (cont.)

Outcome†	Operative Treatment		Nonoperative Treatment		Between-Group Difference or Odds Ratio‡	Significance
	No. of Patients (Missing)	n (%)	No. of Patients (Missing)	n (%)	Odds Ratio (95% CI)	
Secondary outcomes: adverse events§§						
Any adverse event: PP	56 (0)	9 (16.1)	58 (1)	3 (5.2)	0.3 (0.1 to 1.2)	—
Adverse event subtype (PP)						
Surgical site infections		1 (11.1)		0 (0.0)	—	—
Neurologic		3 (33.3)		2 (66.7)	—	—
Thromboembolic		3 (33.3)		0 (0.0)	—	—
Rerupture		1 (11.1)		0 (0.0)	—	—
Other complications		1 (11.1)		1 (33.3)	—	—
OBS COHORT	42 (2)	5 (11.9)	49 (4)	3 (6.1)	0.7 (0.1 to 3.8)	—
Secondary outcomes: return to sports¶¶						
PP						
Full return	54 (2)	31 (57.4)	57 (2)	23 (40.3)	0.5 (0.2 to 1.0)	—
No or partial return		23 (42.6)		34 (69.7)		
OBS COHORT						
Full return	38 (6)	22 (57.9)	45 (8)	20 (44.4)	0.3 (0.1 to 0.7)	—
No or partial return		16 (42.1)		25 (55.6)		

* The as-treated (AS), per-protocol (PP), and observational cohort (OBS COHORT) analyses used propensity scores adjustment by inverse probability weighting. CI denotes confidence interval; IPAQ, International Physical Activity Questionnaire; ITT, intention to treat; LEFS, Lower Extremity Functional Scale; LSI%, limb symmetry index percentage; MET, metabolic equivalent of task; MRI, magnetic resonance imaging; PHAT, Perth Hamstring Assessment Tool; SD, standard deviation; and VAS, visual analog scale.

† PHAT[16] measures patient-reported pain and function during various activities, and scores range from 0 (worst) to 100 (best). The maximum best score of LEFS[18] is 80, and the minimum worst score is 0. The results for PHAT and LEFS were obtained after imputation of internally missing answers. The short form of the IPAQ was used.[19] MET hours were calculated as recommended (https://sites.google.com/view/ipaq/score). Higher MET hours indicate higher activity, with the lowest possible score being 0 and the highest possible score being 321.

‡ Between-group differences were obtained from a linear regression model with adjustment for age, sex, and degree of tendon retraction. Negative numbers are in favor of operative treatment. The widths of the CIs have not been adjusted for multiplicity, and they should not be used to infer treatment effects. Odds ratios were calculated for the adverse events and the ability to return fully to sports.

§ Analysis with multiple imputation by chained equations was used for patients who completely missed the visit in addition to imputation of questions internally missing in PHAT and LEFS.

¶ Patient satisfaction with outcome was assessed on a VAS in which patients were asked to mark on a line to indicate how satisfied they were with their treatment (0=not satisfied at all, 100=fully satisfied).

‖ Patient estimation of recovery was assessed on a VAS in which patients were asked to mark on a line to indicate how well their limb had recovered (0=not recovered at all, 100=fully recovered).

** Functional tests were all performed as described by Pihl et al.[5] LSI was calculated for all tests as: (test injured leg/test uninjured leg) × 100. Measurement of maximum isometric force (Newtons) was performed using a handheld isometric dynamometer. In the single-step test, the number of seconds to perform 20 squats was counted. Single-leg hop and triple hop for distance were measured in centimeters.

†† Total hamstring muscle volume (TMV) is the combined volume of the individually segmented hamstring muscles (semimembranosus, semitendinosus, and the short and long heads of the biceps femoris) on MRI Dixon images. TMV was defined as the sum of all voxels included in a muscle segmentation. TMV LSI% represents the TMV injured leg divided by the TMV uninjured leg × 100.

‡‡ The fat fraction was the total volume of fat on MRI Dixon images within the segmented entire hamstring muscles compartment divided by the total muscle volume. The fat fraction LSI% represents the fat fraction of the injured leg divided by the fat fraction of the uninjured leg × 100.

§§ An adverse event was defined as an untoward medical occurrence after initiation of treatment that was not necessarily caused by the treatment.

¶¶ Return to sports was collected on an ordinal scale (not able to return, partly able to return, and fully able to return). In the analysis, the scale was dichotomized to fully able to return to sports versus not able to return combined with partly able to return.

78.5±19.4 in the nonoperatively treated patients (Fig. 3 and Table S4).

LEFS scores, the IPAQ physical activity estimates, and results on the functional tests, including the isometric muscle strength test, were similar in the randomized treatment groups during follow-up (Table 2 and Table S4).

In the quantitative MRI analysis, the total muscle volume was lower and the fat fraction was higher in the injured

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

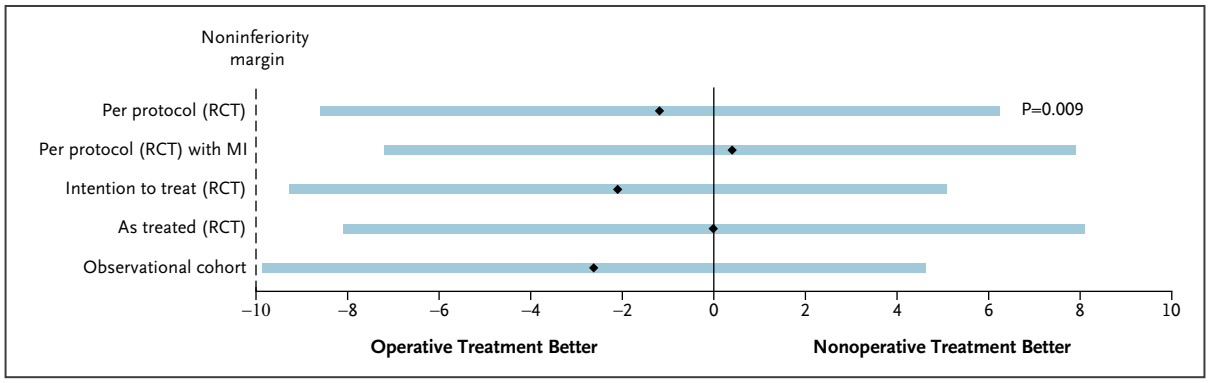


Figure 2. Results of the Primary Outcome.

Per-protocol (with and without multiple imputation [MI]), intention-to-treat, and as-treated results of the primary outcome (difference in total Perth Hamstring Assessment Tool [PHAT] score between operative and nonoperative treatments) are shown for the randomized trial and the observational cohort. PHAT measures patient-reported pain and function during various activities, and scores range from 0 (worst) to 100 (best). Black diamonds indicate the absolute differences in PHAT score between the nonoperative and operative groups, and horizontal blue bars indicate 95% confidence intervals. The per-protocol analysis included randomized participants with follow-up completed at 24 months. We also performed a model-based multiple imputation analysis in which PHAT outcomes for participants who underwent randomization but did not complete the 24-month follow-up visit were imputed based on treatment group, age, sex, and PHAT outcomes at previous follow-up visits. PHAT scores were compared between the two treatment groups using a linear regression model, including an indicator of randomized treatment group assignment, that was adjusted for age, sex, and degree of tendon retraction. If the lower boundary of the two-sided 95% confidence interval for the difference in PHAT score (nonoperative group minus operative group) was greater than −10 points (the dashed vertical line), nonoperative treatment was deemed to be noninferior. P value is for noninferiority of nonoperative treatment compared with operative treatment. In addition, the observational cohort was analyzed using propensity score adjustment by inverse probability weighting. RCT denotes randomized controlled trial.

hamstrings compared with the uninjured hamstrings in both cohorts and regardless of treatment (Table S5). The total muscle volume of the injured hamstrings was somewhat higher in the operative group compared with that in the nonoperative group (LSI percentage, 88 vs. 84; between-group difference, −4.6; 95% CI, −8.6 to −0.6) (Table 2), and operatively treated patients had less fatty infiltration than the nonoperatively treated patients (fat fraction LSI percentage, 137 vs. 152; between-group difference, 20.5; 95% CI, 7.9 to 33.1).

In the randomized trial, more patients in the operative group stated that they had returned fully to sports after 24 months compared with the nonoperative group (57% vs. 40%), yielding an odds ratio and a relative risk for full return to sports of 0.5 (95% CI, 0.2 to 1.0) and 0.7 (95% CI, 0.5 to 1.0), respectively (Table 2 and Tables S6 and S9).

HARMS/ADVERSE EVENTS

In the per-protocol analysis of the randomized trial, 12 adverse events were identified (9 in the operative group vs. 3 in the nonoperative group; odds ratio, 0.3 [95% CI, 0.1 to 1.2]; relative risk, 0.3 [95% CI, 0.1 to 1.3]) (Table 2 and Tables S7 and S9). In the combined randomized and observational cohorts, 20 (9.2%) adverse events were observed (14 [14%] in the operative group and 6 [6%] in the nonoperative group) (Table 2 and Table S7); of these, nine were classified as serious. Six serious adverse events were identified in operatively treated patients and included one myocardial infarction, one deep infection, two deep venous thromboses, and two suspected iatrogenic nerve injuries. In the nonoperatively treated patients, three serious adverse events were identified; two patients had ischialgia treated with neurolysis, and one patient who fell as a result of weakness sustained a concussion. In addition, 101 (47%) patients (59 [58%] in the operative group and 42 [37%] in the nonoperative group) reported persistent neurologic symptoms at 12 and/or 24 months. These symptoms were described in terms of pain, numbness, tingling, cramps, weakness, and, often, a combination of these.

Discussion

We present results challenging the recommendations and corresponding practices that often favor operative over

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

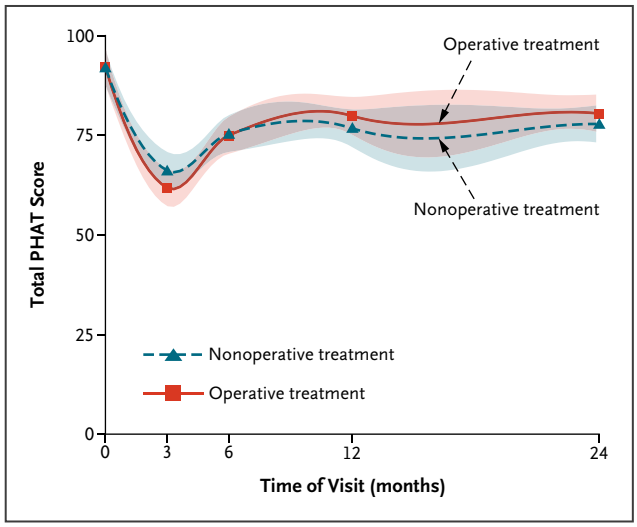


Figure 3. Treatment Outcome Measured by the PHAT Score over Time.

The Perth Hamstring Assessment Tool (PHAT)[16] measures patient-reported pain and function during various activities, and scores range from 0 (worst) to 100 (best). The graph shows the PHAT scores of the operative treatment (red squares) and nonoperative treatment (blue triangles) groups at baseline (preinjury recall) and at 3, 6, 12, and 24 months of follow-up. The mean PHAT score in the randomized trial (the per-protocol population) is represented by the symbols at each time point. Red (solid) and blue (dashed) lines show a smoothed curve fit to the time points with 95% confidence bands.

nonoperative treatment for acute PHA.[6,7,9,25,26] Our randomized trial for patients with PHA shows noninferiority of nonoperative treatment compared with operative treatment. This finding is consistent across various sensitivity analyses and in the examination of the observational cohort.

Our results are concordant with results of recent retrospective and prospective cohort studies.[10-12,27] In a 2022 study by van der Made et al.[12] comparing nonoperative and operative treatment, no significant differences in PHAT scores were observed at 12 months. Likewise, Kanakamedala et al.[10] reported comparable functional results between operative and nonoperative treatment, with a higher incidence of sciatic nerve symptoms in operatively treated patients (53%) versus nonoperatively treated patients (33%).

The analyses of secondary end points reinforce our interpretation of the primary outcome data. At 24 months, the mean between-group difference in LEFS,[28] a validated outcome measure of lower extremity function, was –1.6 (95% CI, –5.2 to 2.0) (Table 2), and there was no difference in any of the functional outcome measures. The follow-up quantitative MRI analyses showed slightly more muscle atrophy and more fat infiltration in nonoperatively treated patients. This could indicate some protective effect of surgery on muscle quality, which in turn, could contribute to the reduction in the probability of returning fully to sports at 24 months of follow-up in the nonoperative group compared with the operative group. Because hamstrings function is of specific importance in high-intensity burst activities, such as sprinting,[29] it is possible that the return-to-sports outcome measure is particularly sensitive to muscle degeneration induced by the injury. However, if this is the case, it did not translate into any observed differences in the specific functional test results, patients' reported activity levels, or overall patient-reported outcomes, nor were there any differences in the proportion of patients not able to return to sports at all at 24 months of follow-up (Table S6). In addition, because surgery inherently carries risks for complications, we unsurprisingly observed more adverse events and a higher proportion of patients reporting persistent neurologic symptoms in the operative group, aligning with findings from other studies.[7,10-12]

For the results of a trial to be clinically useful, they must be applicable to a definable group of patients. However, all randomized trials are subject to selection bias, which can impact their external validity. This bias can arise from stringent inclusion/exclusion criteria, patient preferences, and perceptions of the treatment options' benefits or risks. Often, this bias reflects a lack of equipoise regarding the treatment groups among either the recruiters or the recruited. Assessing the presence and magnitude of selection bias in most randomized trials is difficult, primarily because data on eligible but not recruited patients are not available. In the current trial, we sought to address these concerns by incorporating a parallel observational cohort for patients who did not participate in the randomized trial, monitoring them in an identical manner to randomized participants. This design enables comparisons between the two groups, offering insights that enhance the external validity of our findings. Our approach achieved a high inclusion rate: of 226 eligible patients, only 10 opted out of both the randomized trial and the observational cohort. Despite a somewhat higher dropout ratio in the observational cohort, we attained an overall follow-up rate of 93%. The baseline characteristics between the two cohorts and the results of the primary outcome were similar, suggesting that the results of the randomized trial are generalizable to the nonrandomized patient population.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

However, treatment allocation within the observational cohort was not random. For instance, compared with those who opted for nonoperative treatment, operatively treated patients in this cohort were more likely to be male, younger, and more physically active and have higher tendon retraction. In addition, the ratio of included patients in the randomized versus observational cohorts varied across trial sites, likely due to differences in perceived clinical equipoise among surgeons at different sites. Moreover, it is possible that surgeons' awareness of the observational cohort influenced their recruitment process. This could potentially lead to directing patients who were less active or those with more complex injuries away from the randomized trial. Despite this potential selection bias, the uniformity in PHAT scores and most secondary outcomes across treatment groups in both the randomized trial and the observational cohort underscores the credibility of our primary findings.

The current trial has some additional limitations. First, the sample size of the trial was small, a consequence of the low incidence of avulsions. Second, an evidence-based noninferiority margin for PHAT does not exist, and we decided on a margin empirically. At the time of designing the trial, there was very little information about the clinical importance of an effect size of –10 PHAT points. Recently, we have reported that a small sample of dissatisfied patients scored 15.5 PHAT points lower than satisfied patients several years after PHA.[5] In addition, it seems that the finding of 10 points is between 20 and 28% of the improvement that patients experience during the first year after PHA.[12,16]

Third, we did not perform sham surgery, which may introduce a difference in the placebo effect of the two treatment groups, likely favoring operative treatment, and a lack of blinding of the surgeons and participants.[30] Our design, however, carries the advantage of a more relevant comparison between the treatment groups as nonoperative treatment avoids surgical scars and allows for earlier rehabilitation. Fourth, the pragmatic multicenter design introduces a risk of heterogeneity regarding surgical technique and rehabilitation, which, conversely, may strengthen the trial's generalizability.

Despite the limitations outlined, our trial yields robust results for both primary and secondary outcomes, challenging the widely held practice of operative repair of PHAs. Nonoperative treatment was noninferior to operative treatment, suggesting that it is a viable option for middle-aged patients with acute PHAs.

Disclosures

Supported by Afa Försäkring, Centrum för Idrottsforskning, the Swedish Patient Insurance Company, Uppsala and Östergötalands Counties, the Knut and Alice Wallenberg Foundation, Karolinska Institutet, and the Swedish Research Council.

Author disclosures and other supplementary materials are available at evidence.nejm.org.

Author Affiliations

[1] Department of Clinical Sciences, Danderyds Hospital, Karolinska Institutet, Stockholm
[2] Department of Orthopedics, Danderyds Hospital, Stockholm
[3] Department of Surgical Sciences, Uppsala University, Uppsala University Hospital, Uppsala, Sweden
[4] Institute of Clinical Medicine, University of Oslo, Oslo
[5] Division of Orthopedic Surgery, Oslo University Hospital, Oslo
[6] Department of Orthopedic Surgery, Haukeland University Hospital, Bergen, Norway
[7] Department of Orthopedic Surgery and Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden
[8] Wallenberg Centre for Molecular Medicine, Linköping University, Linköping, Sweden
[9] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm
[10] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm
[11] Department of Orthopaedic Surgery, Østfold Hospital Trust, Grålum, Norway

References

1. Lohmander LS, Roos EM. The evidence base for orthopaedics and sports medicine. BMJ 2015;350:g7835. DOI: 10.1136/bmj.g7835.
2. Lim HC, Adie S, Naylor JM, Harris IA. Randomised trial support for orthopaedic surgical procedures [published correction appears in PLoS One 2015;10:e0144682]. PLoS One 2014;9:e96745. DOI: 10.1371/journal.pone.0096745.
3. Irger M, Willinger L, Lacheta L, Pogorzelski J, Imhoff AB, Feucht MJ. Proximal hamstring tendon avulsion injuries occur predominately in middle-aged patients with distinct gender differences: epidemiologic analysis of 263 surgically treated cases. Knee Surg Sports Traumatol Arthrosc 2020;28:1221-1229. DOI: 10.1007/s00167-019-05717-7.
4. Laszlo S, Jonsson KB. Increasing incidence of surgically treated hamstring injuries: a nationwide registry study in Sweden between 2001 and 2020. Acta Orthop 2023;94:336-341. DOI: 10.2340/17453674.2023.13650.
5. Pihl E, Jonsson KB, Berglöf M, Brodin N, Sköldenberg O, Hedbeck CJ. Exploring the Perth Hamstring Assessment Tool and Lower Extremity Functional Scale in a proximal hamstring avulsion

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.

cohort: a cross-sectional study. Am J Sports Med 2021;49:1732-1740. DOI: 10.1177/03635465211008568.

6. van der Made AD, Reurink G, Gouttebarge V, Tol JL, Kerkhoffs GM. Outcome after surgical repair of proximal hamstring avulsions: a systematic review. Am J Sports Med 2015;43:2841-2851. DOI: 10.1177/0363546514555327.

7. Bodendorfer BM, Curley AJ, Kotler JA, et al. Outcomes after operative and nonoperative treatment of proximal hamstring avulsions: a systematic review and meta-analysis. Am J Sports Med 2018;46:2798-2808. DOI: 10.1177/0363546517732526.

8. Bertiche P, Mohtadi N, Chan D, Hölmich P. Proximal hamstring tendon avulsion: state of the art. J ISAKOS 2021;6:237-246. DOI: 10.1136/jisakos-2019-000420.

9. Laszlo S, Nilsson M, Pihl E, et al. Proximal hamstring tendon avulsions: a survey of orthopaedic surgeons' current practices in the Nordic countries. Sports Med Open 2022;8:49. DOI: 10.1186/s40798-022-00439-6.

10. Kanakamedala AC, Rynecki ND, Mojica ES, et al. No difference in clinical outcomes between operative and nonoperative management of minimally retracted proximal hamstring ruptures. Knee Surg Sports Traumatol Arthrosc 2023;31:2739-2745. DOI: 10.1007/s00167-023-07400-4.

11. Pihl E, Skoldenberg O, Nasell H, Jonhagen S, Kelly Pettersson P, Hedbeck CJ. Patient-reported outcomes after surgical and non-surgical treatment of proximal hamstring avulsions in middle-aged patients. BMJ Open Sport Exerc Med 2019;5:e000511. DOI: 10.1136/bmjsem-2019-000511.

12. van der Made AD, Peters RW, Verheul C, et al. Proximal hamstring tendon avulsions: comparable clinical outcomes of operative and non-operative treatment at 1-year follow-up using a shared decision-making model. Brit J Sports Med 2022;56:340-348. DOI: 10.1136/bjsports-2021-104588.

13. Pihl E, Kristoffersen MH, Rosenlund A-M, et al. The proximal hamstring avulsion clinical trial (PHACT) — a randomised controlled non-inferiority trial of operative versus non-operative treatment of proximal hamstrings avulsions: study protocol. BMJ Open 2019;9:e031607. DOI: 10.1136/bmjopen-2019-031607.

14. Blakeney WG, Zilko SR, Edmonston SJ, Schupp NE, Annear PT. A prospective evaluation of proximal hamstring tendon avulsions: improved functional outcomes following surgical repair. Knee Surg Sports Traumatol Arthrosc 2017;25:1943-1950. DOI: 10.1007/s00167-017-4475-0.

15. Blakeney WG. Editorial commentary: proximal hamstring tendon injuries: is the research hamstrung by the use of poor outcome measures? Arthroscopy 2020;36:1308-1310. DOI: 10.1016/j.arthro.2020.01.015.

16. Blakeney WG, Zilko SR, Edmonston SJ, Schupp NE, Annear PT. Proximal hamstring tendon avulsion surgery: evaluation of the Perth Hamstring Assessment Tool. Knee Surg Sports Traumatol Arthrosc 2017;25:1936-1942. DOI: 10.1007/s00167-016-4214-y.

17. Reza T, Hinkle AJ, Perez-Chaumont A, Brown SM, Mulcahey MK. Systematic review of outcome measures used after proximal hamstring repair. Orthop J Sports Med 2021;9:23259671211005101. DOI: 10.1177/23259671211005101.

18. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. Phys Ther 1999;79:371-383.

19. Craig CL, Marshall AL, Sjöström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35:1381-1395. DOI: 10.1249/01.MSS.0000078924.61453.FB.

20. Marmon AR, McClelland JA, Stevens-Lapsley J, Snyder-Mackler L. Single-step test for unilateral limb ability following total knee arthroplasty. J Orthop Sports Phys Ther 2013;43:66-73. DOI: 10.2519/jospt.2013.4372.

21. Noyes FR, Barber SD, Mangine RE. Abnormal lower limb symmetry determined by function hop tests after anterior cruciate ligament rupture. Am J Sports Med 1991;19:513-518. DOI: 10.1177/036354659101900518.

22. Reid A, Birmingham TB, Stratford PW, Alcock GK, Giffin JR. Hop testing provides a reliable and valid outcome measure during rehabilitation after anterior cruciate ligament reconstruction. Phys Ther 2007;87:337-349. DOI: 10.2522/ptj.20060143.

23. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine — reporting of subgroup analyses in clinical trials. N Engl J Med 2007;357:2189-2194. DOI: 10.1056/NEJMsr077003.

24. van der Made AD, Smithuis FF, Buckens CF, et al. Good interrater reliability for standardized MRI assessment of tendon discontinuity and tendon retraction in acute proximal full-thickness hamstring tendon injury. Am J Sports Med 2021;49:2475-2481. DOI: 10.1177/03635465211021612.

25. Hillier-Smith R, Paton B. Outcomes following surgical management of proximal hamstring tendon avulsions : a systematic review and meta-analysis. Bone Jt Open 2022;3:415-422. DOI: 10.1302/2633-1462.35.BJO-2021-0196.R1.

26. Pasic N, Giffin JR, Degen RM. Practice patterns for the treatment of acute proximal hamstring ruptures. Phys Sportsmed 2020;48:116-122. DOI: 10.1080/00913847.2019.1645576.

27. Shambaugh BC, Olsen JR, Lacerte E, Kellum E, Miller SL. A comparison of nonoperative and operative treatment of complete proximal hamstring ruptures. Orthop J Sports Med 2017;5:2325967117738551. DOI: 10.1177/2325967117738551.

28. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement properties of the Lower Extremity Functional Scale: a systematic review. J Orthop Sports Phys Ther 2016;46:200-216. DOI: 10.2519/jospt.2016.6165.

29. Kenneally-Dabrowski CJB, Brown NAT, Lai AKM, Perriman D, Spratford W, Serpell BG. Late swing or early stance? A narrative review of hamstring injury mechanisms during high-speed running. Scand J Med Sci Sports 2019;29:1083-1091. DOI: 10.1111/sms.13437.

30. Dettori JR, Norvell DC, Chapman JR. The art of surgery: the strange world of the placebo response. Global Spine J 2019;9:680-683. DOI: 10.1177/2192568219861972.

For personal use only. No other uses without permission. Copyright © 2024 Massachusetts Medical Society.