ADVANCING THE MANAGEMENT OF PATELLAR TENDINOPATHY

Improving patient education and understanding treatment response

Jie Deng
邓婕

ADVANCING THE MANAGEMENT OF PATELLAR TENDINOPATHY

improving patient education and understanding treatment response

Jie Deng

邓婕

ISBN/EAN: 978-94-6534-093-7

Cover Design: Jie Deng, Room of Rebellion Design Studio
Layout and Printed by: ProefschriftMaken

Publication of this thesis was financially supported by the Department of Orthopedics and Sports Medicine, Department of Radiology and Nuclear Medicine at the Erasmus MC, the Erasmus University Rotterdam.

ADVANCING THE MANAGEMENT OF PATELLAR TENDINOPATHY

improving patient education and understanding treatment response

Optimalisatie van de behandeling van patella tendinopathie

Verbetering van patiënt educatie en inzicht in behandelrespons

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr.ir. A.J. Schuit
and in accordance with the decision of the Doctorate Board.
The public defence shall be held on
Tuesday 27 January 2026 at 15.30 hrs

by

Jie Deng
born in Sichuan, People's Republic of China.

Erasmus
University
Rotterdam

Doctoral Committee:

Promotors:	Prof.dr. E.H.G. Oei
	Prof.dr. D. Eygendaal
Other members:	Dr. M. van Middelkoop
	Prof.dr. L.M.G. Hartgens
	Prof.dr. P. Van Dyck
Co-promotor:	Dr. R.J. de Vos

Table of Content

Abbreviations

AI: Artificial Intelligence

AP: Anteroposterior

CSAS: Cincinnati Sports Activity Scale

DAGs: Directed Acyclic Graphs

EET: Eccentric Exercise Therapy

EQ-5D-3L: European Quality of Life–3 Dimensions

FKGL: Flesch-Kincaid Grade Level

GROC: Global Rating of Change

GSUS: Grey-scale Ultrasound

IFP: Infrapatellar Fat Pad

IQR: Interquartile Range

LLM: Large Language Model

LRT: Likelihood Ratio Statistic

MCID: Minimal Clinically Important Difference

MRI: Magnetic Resonance Imaging

NDE: Natural Direct Effect

NIE: Natural Indirect Effect

PDUS: Power Doppler Ultrasound

PEMAT: Patient Education Materials Assessment Tool

PT: Patellar Tendinopathy

PTLE: Progressive Tendon Loading Exercises

SD: Standard Deviation

STS: Semantic Textual Similarity

TE: Total Effect

US: Ultrasound

VAS-SLDS: Pain after Single-leg Decline Squat

VISA-P: Victorian Institute of Sports Assessment-Patella

1

Chapter 1

General Introduction

Patellar tendinopathy

Patellar tendinopathy (PT), also called “Jumper’s knee”, is the preferred term for a condition characterized by localized pain at the patellar tendon and reduced functional capacity affecting both work productivity and sports performance.[1,2] PT is frequently diagnosed in athletes involved in repetitive jumping sports, with a prevalence of 20.8% in basketball players and 24.8% in volleyball players.[3] Although PT can occur during adolescence, it is more commonly observed in adult athletes.[2,3] The prevalence of PT increases with age, especially those between 15 and 30 years old, and is twice as high in adult athletes as in the younger population.[3,13] This disorder considerably impacts athletes, with 58% reporting reduced work productivity and 55% reducing or ceasing their preferred sports activities.[4]

PT is a multifactorial condition, and several risk factors are involved in developing PT.[2,5,6] Male sex is suggested as a non-modifiable risk factor that may increase the development of PT.[6] Several modifiable risk factors have also been described, which could help clinicians design prevention programs for PT by addressing these factors.[5] A systematic review identified decreased ankle dorsiflexion, decreased flexibility of the quadriceps and the hamstrings as risk factors for the development of PT, though evidence from included prospective studies was limited.[5] The presence of ultrasound abnormalities at the patellar tendon as hypo-echogenicity, tendon thickening or neovascularity, can also be an intrinsic risk factor, as it has been associated with future symptoms at the patellar tendon in a prospective study.[2,7] Other extrinsic factors found in prospective studies, such as increased training volume, greater jumping height, participation in certain sports and occupation, have also been considered as risk factors for PT.[5,6]

Histopathological examination of PT has shown degenerative changes in the extracellular matrix (ECM), including increased levels of type III collagen and glycosaminoglycan (GAG) accumulation.[8,9] At the cellular level, changes include increased cellularity, the presence of inflammatory mediators and immune cells, and morphological alterations in tenocytes.[2,10] In addition, neovascularization or abnormal vascular proliferation is also observed.[11] These features can lead to an increase in tendon thickness and an altered fibrillar pattern, commonly identified on US or MRI as characteristic findings of PT. (Figure 1).[12]

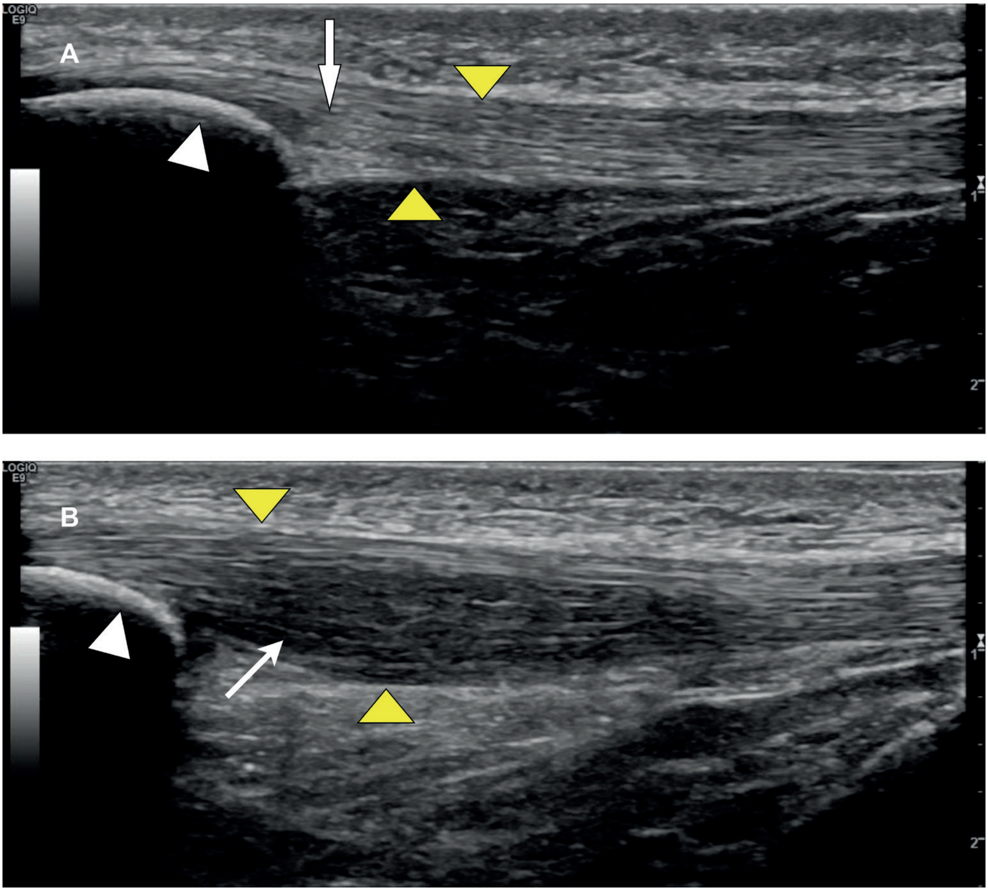


Figure 1: Longitudinal grey-scale ultrasound of normal patellar tendon (A) and patellar tendinopathy (B)

A: Grey-scale ultrasound image of an asymptomatic athlete with a normal patellar tendon (arrow) appears hyperechoic with a fiber-like or linear fibrillar echotexture arising from the inferior pole of the patella (solid arrowhead). The tendon shows well-defined echogenic margins anteriorly and posteriorly (yellow solid arrowhead).

B: Grey-scale ultrasound image of a basketball player with PT shows marked thickening of proximal patellar tendon (yellow solid arrowhead). Note a focal area of low echogenicity in posterior portion of tendon (arrow).

The clinical diagnosis of PT is based on the patient's history of activity-provoked localized tendon pain (Figure 2) and physical examination with pain on loading tests and recognizable pain on tendon palpation. Loading tests that have been suggested as pain-provoking tests are key to assisting in confirming the diagnosis, such as single-leg decline squat (SLDS).[2,13] Imaging focusing on altered tendon morphology (e.g., thickening) or structure (e.g., hypo-echogenicity) may be used to support the clinical diagnosis when uncertainty exists or when symptoms deviate from the expected course during follow-up.[14] Conventional ultrasound (US), including grey-scale and power Doppler ultrasound, remains recommended as the first imaging modality of choice because of its wide availability and low costs. If necessary, MRI can be considered to rule out coexisting pathologies, such as cartilage injuries.[2,13,15]

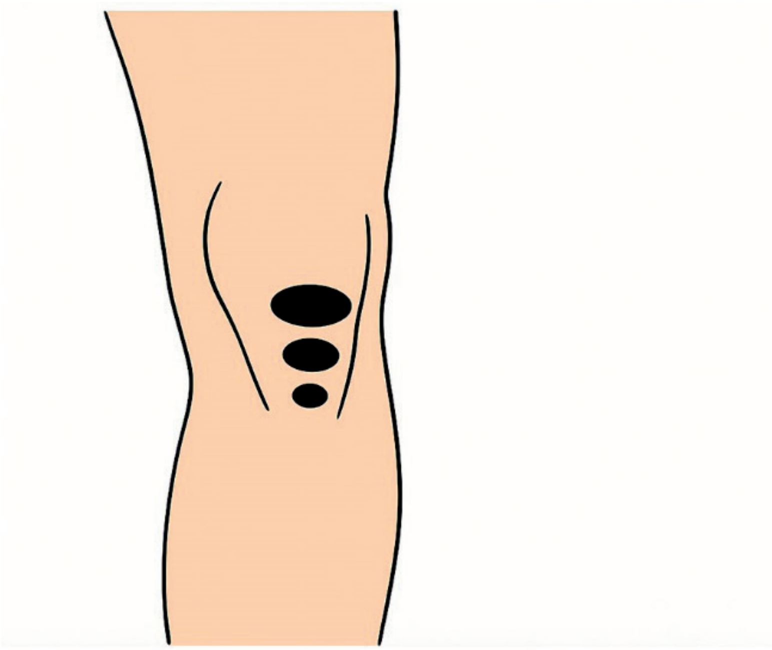


Figure 2: Localized pain at the patellar tendon

Current PT management recommendations include comprehensive patient education and exercise-based interventions, with loading programs being the first choice.[1,16] Over the past two decades, eccentric exercise has been the primary treatment for PT.[1,2] The benefits from this heavy-load modality may be linked to generating greater mechanical stimulation of the patellar tendon, allowing for the adaptation of the musculotendinous unit and restoring tendon capacity.[1,2] However, this program is pain-provoking, and lacks effectiveness when applied during the sports season.[13,17] In contrast, a progressive exercise therapy regimen is designed to gradually reduce pain and restore tendon function.[13,18] More recently, the largest randomized trial in athletes with PT to date (JUMPER study), conducted in our group, demonstrated that progressive tendon loading exercises (PTLE) reduced disability at 24 weeks compared with eccentric exercises.[18] These results have changed the Dutch guideline and shown potential as the initial conservative treatment for PT.[1,19]

Adjunct treatments such as platelet-rich plasma, hyaluronic acid injection, or shockwave therapy are available for PT and may be considered if exercise and education are not sufficiently effective.[1,2,20] However, their effectiveness alone or combined with exercises in PT remains debatable and requires further research.[1,20,21] Surgery (e.g., patellar tendon debridement) might be an option in refractory cases after conservative management fails, but a limited number of randomized trials comparing surgery with exercise treatment and potential complications warrant caution when considering invasive approaches.[22,23]

The long-term prognosis of PT is variable, with many patients failing to show full resolution of symptoms, highlighting the need for more effective, targeted, patient-centered management strategies.[24] Therefore, this thesis focuses on improving the management of PT by enhancing patient education and investigating whether clinically measurable factors (physical tests and conventional imaging modalities) can be used to understand treatment response in patients with PT undergoing exercise-based treatment.

Patient education

Patient education is a vital component of PT management.[2] Education on the condition, a realistic recovery timeline, and addressing psychosocial factors are critical for promoting adherence to rehabilitation programs and self-management.[2,15] Ultimately, such a comprehensive patient education scheme may increase the success of the prescribed loading protocols and forge a therapeutic alliance between clinicians and patients.[2,15] For decades, face-to-face communication has remained the primary format for clinicians to communicate educational knowledge when managing individuals with chronic musculoskeletal conditions.[25,26] However, this approach can be time-consuming, resource-intensive, and unsustainable. Educating patients via telehealth (e.g., video conferencing or telephone) may help overcome geographic barriers, but it is not yet widely adopted in routine clinical practice.[26-31] These clinician-delivered education approaches still struggle to provide timely advice or support that is often needed when individuals with PT experience frustration due to slow progress or discomfort during loading-based treatment.

The Internet has become the most frequent primary source of medical information for the general population.[32] People with chronic musculoskeletal conditions, such as tendinopathy, are no exception to this trend. More than 50% of orthopedic patients went looking for their condition online before consulting a doctor, with younger patients being more likely to do so[33]. A large number of patients perceive online orthopedic information as useful for promoting self-management of their chronic disease in an easily accessible way.[32-35] Despite these benefits, a bulk of concerns have been identified regarding web-based patient education materials for musculoskeletal conditions, including a lack of quality and personalized information, as well as being incompatible with individuals' health literacy.[35-38] These obstacles may be addressed by large language models (LLMs), a form of advanced artificial intelligence (AI) capable of "understanding" natural language input and generating human-like responses.[39] Trained on a vast amount of online content, including biomedical literature and current guidelines, LLMs may augment patient education by producing evidence-based information to assist patients in communicating with healthcare providers in a digestible, on-demand, and interactive manner.[39,40] To date, one widely used LLM-driven chatbot is ChatGPT, which allows users to submit queries or prompts and generate responses based on its contextual understanding of the input.[41] Recent research has demonstrated that ChatGPT can generate accurate responses to questions related to musculoskeletal disease, including osteoarthritis[41] and anterior cruciate ligament (ACL) surgery.[42] These results suggest that this chatbot has the potential to serve as an accessible educational tool. However, ChatGPT's performance in answering PT-related questions is currently unexplored.

Two previous cohort studies have shown that more than half of PT athletes quit sports after 15 years, and 19% retired from competitive sports after 11 years due to persistent knee symptoms.[24,43] However, these studies only included small sample sizes, and participants did not receive the current standardized exercise-based strategies. Providing updated education on recovery time frames, often referred to as prognostic information, is urgently needed, as questions about prognosis are among the most important for patients with PT undergoing rehabilitation.[44] Knowledge on prognosis for PT is primarily based on studies conducted over two decades ago or is limited to specific populations, such as elite volleyball players or men, making it difficult to set realistic goals for PT patients.[24,43]

Understanding treatment response using clinical and imaging parameters

Response to exercise-based strategies can be quantified by changes in pain intensity (measured on a 0–10 numeric rating scale) and pain-related disability (measured using the Victorian Institute of Sports Assessment–Patellar [VISA-P]). As core domains in PT, these outcome measures are capable of detecting clinically meaningful change.[45-47] Besides these patient-reported outcomes, the function of the kinetic chain is assessed before treatment in practice and clinical research.[13,15,18,48] A thorough examination of the kinetic chain can be objectively assessed with physical tests.[13] Reduced strength in quadriceps and hip abductors[49], limited ankle dorsiflexion range of motion, and decreased flexibility in quadriceps and hamstrings[50] have been commonly observed in patients with PT. Jumping performance can also be assessed to assess energy-storage and release ability.[13,49,50] However, the evidence is conflicting, with some authors reporting that vertical jumping height is greater in subjects with PT compared to asymptomatic individuals.[5] Conventional US can be used to assess tendon morphology (e.g., tendon thickness) and the presence of abnormalities (e.g., intratendinous calcification, bone erosions at the inferior pole of the patella, or neovascularization).[51] MRI may further identify abnormalities in adjacent structures, such as the presence of patella bone marrow edema and infrapatellar fat pad (IFP, or Hoffa's fat pad) edema.[52]

Our previous trial has demonstrated a moderate but significant difference in VISA-P score at 24 weeks between PTLE over eccentric exercise[18], warranting further investigation into the underlying mechanism by which PTLE may work by identifying potential targets or mediators that could optimize treatment outcomes.[18,53] Since loading-based exercise may normalize muscle performance or structural abnormalities in the patellar tendon, it is reasonable to assume that PTLE may exert its effects by targeting these factors. No study has investigated this hypothesis within a causal mediation framework.[54,55]

The response to exercise treatment among individuals also varies, with some patients showing resolution in a few weeks and others improving in months.[18,48] One strategy to enhance patient outcomes is to predict who is at risk of poor outcomes by identifying prognostic factors (Figure 3)[56]. While most functional deficits assessed by physical tests or US-detected features could serve as risk factors and inform prevention strategies, it is currently unknown whether they have prognostic value for predicting treatment outcomes. Besides, the presence of infrapatellar fat pad (IFP, or Hoffa's fat pad) edema or/and patellar bone edema on MRI were associated with poor recovery in individuals with PT undergoing arthroscopic patellar release[52]. Their prognostic value for predicting future clinical outcomes following exercise-based treatment also remains unclear.

To address these gaps, there is a need to clarify the role of clinical and imaging assessments in understanding the treatment response of exercise-based interventions for PT.

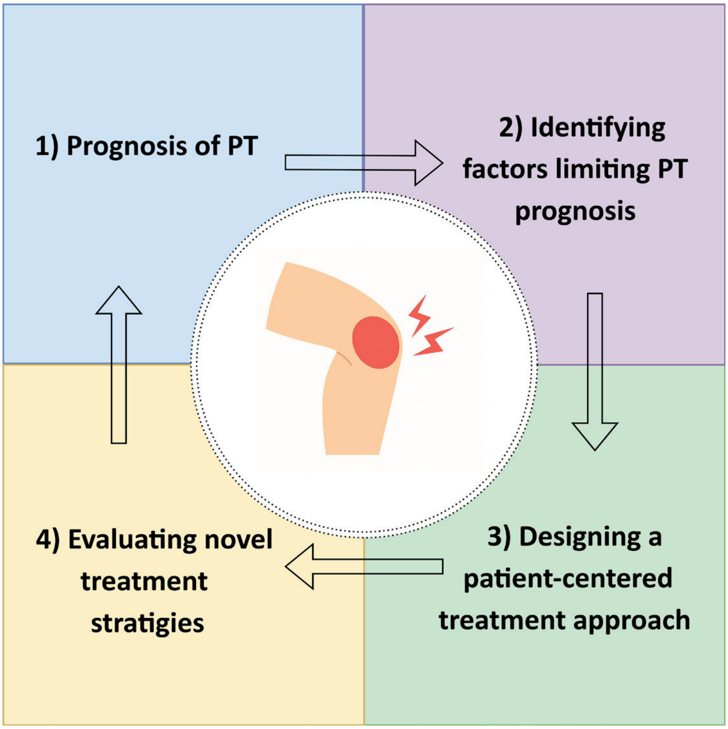


Figure 3: Illustration of the role of identifying prognostic factors in clinical research and practice

Aims and outline

The thesis consists of two parts. In **Part I**, we aimed to enhance patient education for PT patients, and in **Part II,** we explored factors to understand treatment response of exercise-based interventions for PT.

Part I consists of **Chapters 2 and 3**. In **Chapter 2,** we collect a large number of patient-centered questions from varied sources and evaluat the output from ChatGPT-4 using a multilevel assessment. This chapter provides fundamental information on its performance in answering PT-related questions and explores whether this chatbot can be used as a potential patient education tool. In **Chapter 3**, we design a follow-up digital questionnaire and revisit 76 PT athletes from the JUMPER study (to our knowledge the largest reported trial in PT) after 5 years. Eventually, we report the overall prognosis using patient-reported outcomes. **Part II** includes **Chapters 4 to 6.** In **Chapter 4**, we conduct causal mediation analyses to investigate whether physical or sonographic factors mediate the superior effectiveness of progressive tendon loading exercises (PTLE) compared to eccentric exercise therapy (EET). **Chapter 5** studies the association between baseline physical test results and the 24-week change in pain-related disability. The association between baseline findings on sonography or MRI with the 24-week change in pain-related disability is explored using mixed-effect models in **Chapter 6**. A general discussion and summary are provided in **Chapters 7 and 8** of this thesis, respectively.

References

1. Marigi EM, Buckley P, Razi F, Abbas MJ, Jildeh TR, Camp CL, Krych AJ, Okoroha KR. Patellar Tendinopathy: Critical Analysis Review of Current Nonoperative Treatments. *JBJS Rev*. 2022;10(3) http://doi.org/e21.00168.
2. Millar NL, Silbernagel KG, Thorborg K, Kirwan PD, Galatz LM, Abrams GD, Murrell GAC, McInnes IB, Rodeo SA. Tendinopathy. *Nat Rev Dis Primers*. 2021/01/07 2021;7(1):1. http://doi.org/10.1038/s41572-020-00234-1.
3. Nutarelli S, da Lodi CMT, Cook JL, Deabate L, Filardo G. Epidemiology of Patellar Tendinopathy in Athletes and the General Population: A Systematic Review and Meta-analysis. *Orthopaedic Journal of Sports Medicine*. 2023;11(6):23259671231173659. http://doi.org/10.1177/23259671231173659.
4. De Vries AJ, Koolhaas W, Zwerver J, Diercks RL, Nieuwenhuis K, Van Der Worp H, Brouwer S, Van Den Akker-Scheek I. The impact of patellar tendinopathy on sports and work performance in active athletes. *Res Sports Med*. 2017;25(3):253-265. http://doi.org/10.1080/15438627.2017.1314292.
5. Sprague AL, Smith AH, Knox P, Pohlig RT, Grävare Silbernagel K. Modifiable risk factors for patellar tendinopathy in athletes: a systematic review and meta-analysis. *Br J Sports Med*. 2018;52(24):1575-1585. http://doi.org/10.1136/bjsports-2017-099000.
6. de Vries AJ, van der Worp H, Diercks RL, van den Akker-Scheek I, Zwerver J. Risk factors for patellar tendinopathy in volleyball and basketball players: A survey-based prospective cohort study. *Scand J Med Sci Sports*. 2015;25(5):678-684. http://doi.org/10.1111/sms.12294.
7. Cushman DM, Stokes D, Vu L, Corcoran B, Fredericson M, Eby SF, Teramoto M. Ultrasound as a predictor of time-loss injury for the patellar tendon, Achilles tendon and plantar fascia in division I collegiate athletes. *Br J Sports Med*. 2025;59(4):241-248. http://doi.org/10.1136/bjsports-2024-109066.
8. Riley G. The pathogenesis of tendinopathy. A molecular perspective. *Rheumatology (Oxford)*. 2004;43(2):131-142.
9. Samiric T, Parkinson J, Ilic MZ, Cook J, Feller JA, Handley CJ. Changes in the composition of the extracellular matrix in patellar tendinopathy. *Matrix Biol*. 2009;28(4):230-236. http://doi.org/10.1016/j.matbio.2009.04.001.
10. Xu Y, Murrell GA. The basic science of tendinopathy. *Clin Orthop Relat Res*. Jul 2008;466(7):1528-38.
11. Pufe T, Petersen WJ, Mentlein R, Tillmann BN. The role of vasculature and angiogenesis for the pathogenesis of degenerative tendons disease. *Scand J Med Sci Sports*. 2005;15(4):211-222.
12. Warden SJ, Kiss ZS, Malara FA, Ooi ABT, Cook JL, Crossley KM. Comparative accuracy of magnetic resonance imaging and ultrasonography in confirming clinically diagnosed patellar tendinopathy. *Am J Sports Med*. 2007;35(3):427-436.
13. Malliaras P, Cook J, Purdam C, Rio E. Patellar Tendinopathy: Clinical Diagnosis, Load Management, and Advice for Challenging Case Presentations. *J Orthop Sports Phys Ther*. 2015;45(11):887-898. http://doi.org/10.2519/jospt.2015.5987.
14. de Vos R-J, van der Vlist AC, Zwerver J, Meuffels DE, Smithuis F, van Ingen R, van der Giesen F, Visser E, Balemans A, Pols M, Veen N, den Ouden M, Weir A. Dutch multidisciplinary guideline on Achilles tendinopathy. *Br J Sports Med*. 2021;55(20):1125-1134. http://doi.org/10.1136/bjsports-2020-103867.
15. Rosen AB, Wellsandt E, Nicola M, Tao MA. Clinical Management of Patellar Tendinopathy. *Journal of Athletic Training*. 2022;57(7):621-631. http://doi.org/10.4085/1062-6050-0049.21.

16. Theodorou A, Komnos G, Hantes M. Patellar tendinopathy: an overview of prevalence, risk factors, screening, diagnosis, treatment and prevention. *Arch Orthop Trauma Surg*. 2023;143(11):6695-6705. http://doi.org/10.1007/s00402-023-04998-5.
17. Visnes H, Hoksrud A, Cook J, Bahr R. No effect of eccentric training on jumper's knee in volleyball players during the competitive season: a randomized clinical trial. *Clin J Sport Med*. 2005;15(4):227-234.
18. Breda SJ, Oei EHG, Zwerver J, Visser E, Waarsing E, Krestin GP, de Vos R-J. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. *Br J Sports Med*. 2021;55(9):501-509. http://doi.org/10.1136/bjsports-2020-103403.
19. Ophey M, Koëter S, van Ooijen L, van Ark M, Boots F, Ilbrink S, Lankhorst NA, Piscaer T, Vestering M, den Ouden Vierwind M, van Linschoten R, van Berkel S. Dutch multidisciplinary guideline on anterior knee pain: Patellofemoral pain and patellar tendinopathy. *Knee Surg Sports Traumatol Arthrosc*. 2025;33(2):457-469. http://doi.org/10.1002/ksa.12367.
20. Challoumas D, Pedret C, Biddle M, Ng NYB, Kirwan P, Cooper B, Nicholas P, Wilson S, Clifford C, Millar NL. Management of patellar tendinopathy: a systematic review and network meta-analysis of randomised studies. *BMJ Open Sport Exerc Med*. 2021;7(4):e001110. http://doi.org/10.1136/bmjsem-2021-001110.
21. Scott A, LaPrade RF, Harmon KG, Filardo G, Kon E, Della Villa S, Bahr R, Moksnes H, Torgalsen T, Lee J, Dragoo JL, Engebretsen L. Platelet-Rich Plasma for Patellar Tendinopathy: A Randomized Controlled Trial of Leukocyte-Rich PRP or Leukocyte-Poor PRP Versus Saline. *Am J Sports Med*. 2019;47(7):1654-1661. http://doi.org/10.1177/0363546519837954.
22. Bahr R, Fossan B, Løken S, Engebretsen L. Surgical treatment compared with eccentric training for patellar tendinopathy (Jumper's Knee). A randomized, controlled trial. *The Journal of Bone and Joint Surgery American Volume*. 2006;88(8):1689-1698.
23. Arner JW, Kaeding CC, Bradley JP. Management of Patellar Tendinopathy. *Arthroscopy*. 2024;40(1):13-15. http://doi.org/10.1016/j.arthro.2023.09.004.
24. Kettunen JA, Kvist M, Alanen E, Kujala UM. Long-term prognosis for jumper's knee in male athletes. A prospective follow-up study. *Am J Sports Med*. 2002;30(5):689-692.
25. (UK) NGC. *Evidence review for communication between healthcare professionals and people with chronic pain (chronic primary pain and chronic secondary pain)*. vol No. 193. NICE Guideline. National Institute for Health and Care Excellence (NICE); 2021.
26. Malliaras P, Merolli M, Williams CM, Caneiro JP, Haines T, Barton C. 'It's not hands-on therapy, so it's very limited': Telehealth use and views among allied health clinicians during the coronavirus pandemic. *Musculoskelet Sci Pract*. 2021;52:102340. http://doi.org/10.1016/j.msksp.2021.102340.
27. Buvik A, Bergmo TS, Bugge E, Smaabrekke A, Wilsgaard T, Olsen JA. Cost-Effectiveness of Telemedicine in Remote Orthopedic Consultations: Randomized Controlled Trial. *J Med Internet Res*. 2019;21(2):e11330. http://doi.org/10.2196/11330.
28. Cottrell MA, Russell TG. Telehealth for musculoskeletal physiotherapy. *Musculoskelet Sci Pract*. 2020;48:102193. http://doi.org/10.1016/j.msksp.2020.102193.
29. Shariat A, Hajialiasgari F, Alizadeh A, Anastasio AT. The role of telehealth in the care of musculoskeletal pain conditions after COVID-19. *Work*. 2023;74(4):1261-1264. http://doi.org/10.3233/wor-220524.
30. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. *Pain*. 2022;163(2):e165-e181. http://doi.org/10.1097/j.pain.0000000000002364.

31. Barton CJ, Ezzat AM, Merolli M, Williams CM, Haines T, Mehta N, Malliaras P. "It's second best": A mixed-methods evaluation of the experiences and attitudes of people with musculoskeletal pain towards physiotherapist delivered telehealth during the COVID-19 pandemic. *Musculoskelet Sci Pract*. 2022;58:102500. http://doi.org/10.1016/j.msksp.2021.102500.
32. Swoboda CM, Van Hulle JM, McAlearney AS, Huerta TR. Odds of talking to healthcare providers as the initial source of healthcare information: updated cross-sectional results from the Health Information National Trends Survey (HINTS). *BMC Fam Pract*. 2018;19(1):146. http://doi.org/10.1186/s12875-018-0805-7.
33. Burrus MT, Werner BC, Starman JS, Kurkis GM, Pierre JM, Diduch DR, Hart JM. Patient Perceptions and Current Trends in Internet Use by Orthopedic Outpatients. *HSS J*. 2017;13(3):271-275. http://doi.org/10.1007/s11420-017-9568-2.
34. Koenig S, Nadarajah V, Smuda MP, Meredith S, Packer JD, Henn RF. Patients' Use and Perception of Internet-Based Orthopaedic Sports Medicine Resources. *Orthopaedic Journal of Sports Medicine*. 2018;6(9):2325967118796469. http://doi.org/10.1177/2325967118796469.
35. Win KT, Hassan NM, Oinas-Kukkonen H, Probst Y. Online Patient Education for Chronic Disease Management: Consumer Perspectives. *J Med Syst*. 2016;40(4):88. http://doi.org/10.1007/s10916-016-0438-0.
36. Schwarz I, Houck DA, Belk JW, Hop J, Bravman JT, McCarty E. The Quality and Content of Internet-Based Information on Orthopaedic Sports Medicine Requires Improvement: A Systematic Review. *Arthrosc Sports Med Rehabil*. 2021;3(5):e1547-e1555. http://doi.org/10.1016/j.asmr.2021.05.007.
37. Badarudeen S, Sabharwal S. Assessing readability of patient education materials: current role in orthopaedics. *Clin Orthop Relat Res*. 2010;468(10):2572-2580. http://doi.org/10.1007/s11999-010-1380-y.
38. Roberts H, Zhang D, Dyer GSM. The Readability of AAOS Patient Education Materials: Evaluating the Progress Since 2008. *The Journal of Bone and Joint Surgery American Volume*. 2016;98(17):e70. http://doi.org/10.2106/jbjs.15.00658.
39. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. 2023;29(8):1930-1940. http://doi.org/10.1038/s41591-023-02448-8.
40. Will ChatGPT transform healthcare? *Nat Med*. 2023;29(3):505-506. http://doi.org/10.1038/s41591-023-02289-5.
41. Cao M, Wang Q, Zhang X, Lang Z, Qiu J, Yung PS-H, Ong MT-Y. Large language models' performances regarding common patient questions about osteoarthritis: A comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and perplexity. *J Sport Health Sci*. 2024:101016. http://doi.org/10.1016/j.jshs.2024.101016.
42. Kaarre J, Feldt R, Keeling LE, Dadoo S, Zsidai B, Hughes JD, Samuelsson K, Musahl V. Exploring the potential of ChatGPT as a supplementary tool for providing orthopaedic information. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(11):5190-5198. http://doi.org/10.1007/s00167-023-07529-2.
43. Visnes H, Bache-Mathiesen LK, Yamaguchi T, Gilhuus HP, Algaard KRH, Hisdal E, Bahr R. Long-term Prognosis of Patellar Tendinopathy (Jumper's Knee) in Young, Elite Volleyball Players: Tendon Changes 11 Years After Baseline. *Am J Sports Med*. 2024:3635465241284648. http://doi.org/10.1177/03635465241284648.

44. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KGM, Steyerberg EW, Roberts I, Schroter S, Altman DG, Riley RD. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ*. 2013;346:e5595. http://doi.org/10.1136/bmj.e5595.
45. Vicenzino B, de Vos R-J, Alfredson H, Bahr R, Cook JL, Coombes BK, Fu SN, Gravare Silbernagel K, Grimaldi A, Lewis JS, Maffulli N, Magnusson SP, Malliaras P, Mc Auliffe S, Oei EHG, Purdam C, Rees JD, Rio EK, Scott A, Speed C, Akker-Scheek Ivd, Weir A, Wolf JM, Zwerver J. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. *Br J Sports Med*. 2020;54(8):444-451. http://doi.org/10.1136/bjsports-2019-100894.
46. Hernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. Mar 2014;48(6):453-7.
47. Challoumas D, Zouvani A, Creavin K, Murray E, Crosbie G, Ng N, Millar NL. Determining minimal important differences for patient-reported outcome measures in shoulder, lateral elbow, patellar and Achilles tendinopathies using distribution-based methods. *BMC Musculoskelet Disord*. 2023;24(1):158. http://doi.org/10.1186/s12891-023-06261-9.
48. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, Couppé C, Hjortshoej MH, Doessing S, Kjaer M, Magnusson SP. Clinical Outcomes, Structure, and Function Improve With Both Heavy and Moderate Loads in the Treatment of Patellar Tendinopathy: A Randomized Clinical Trial. *Am J Sports Med*. 2021;49(4):982-993. http://doi.org/10.1177/0363546520988741.
49. Mendonça LD, Ocarino JM, Bittencourt NFN, Macedo LG, Fonseca ST. Association of Hip and Foot Factors With Patellar Tendinopathy (Jumper's Knee) in Athletes. *J Orthop Sports Phys Ther*. 2018;48(9):676-684. http://doi.org/10.2519/jospt.2018.7426.
50. van der Worp H, van Ark M, Roerink S, Pepping G-J, van den Akker-Scheek I, Zwerver J. Risk factors for patellar tendinopathy: a systematic review of the literature. *Br J Sports Med*. 2011;45(5):446-452. http://doi.org/10.1136/bjsm.2011.084079.
51. Carr JC, Hanly S, Griffin J, Gibney R. Sonography of the patellar tendon and adjacent structures in pediatric and adult patients. *AJR Am J Roentgenol*. 2001;176(6):1535-1539.
52. Ogon P, Izadpanah K, Eberbach H, Lang G, Südkamp NP, Maier D. Prognostic value of MRI in arthroscopic treatment of chronic patellar tendinopathy: a prospective cohort study. *BMC Musculoskelet Disord*. 2017;18(1):146. http://doi.org/10.1186/s12891-017-1508-2.
53. Cashin AG, McAuley JH, VanderWeele TJ, Lee H. Understanding how health interventions or exposures produce their effects using mediation analysis. *BMJ*. 2023;382:e071757. http://doi.org/10.1136/bmj-2022-071757.
54. Malliaras P, Barton CJ, Reeves ND, Langberg H. Achilles and patellar tendinopathy loading programmes: a systematic review comparing clinical outcomes and identifying potential mechanisms for effectiveness. *Sports Med*. 2013;43(4):267-286. http://doi.org/10.1007/s40279-013-0019-z.
55. Rabello LM, van den Akker-Scheek I, Brink MS, Maas M, Diercks RL, Zwerver J. Association Between Clinical and Imaging Outcomes After Therapeutic Loading Exercise in Patients Diagnosed With Achilles or Patellar Tendinopathy at Short- and Long-Term Follow-up: A Systematic Review. *Clin J Sport Med*. 2020;30(4):390-403. http://doi.org/10.1097/jsm.0000000000000624.
56. Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, Malats N, Briggs A, Schroter S, Altman DG, Hemingway H. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med*. 2013;10(2):e1001380. http://doi.org/10.1371/journal.pmed.1001380.

2

Chapter 2

ChatGPT Is a Comprehensive Education Tool for Patients With Patellar Tendinopathy, but It Currently Lacks Accuracy and Readability

Jie Deng, Lun Li, Jelle J Oosterhof, Peter Malliaras, Karin Grävare Silbernagel, Stephan J Breda, Denise Eygendaal, Edwin HG Oei, Robert-Jan de Vos

Musculoskelet Sci and Pract. 2025 Apr:76:103275.

Abstract

Background: Generative artificial intelligence tools, such as ChatGPT, are becoming increasingly integrated into daily life, and patients might turn to this tool to seek medical information.

Objective: To evaluate the performance of ChatGPT-4 in responding to patient-centered queries for patellar tendinopathy (PT).

Methods: Forty-eight patient-centered queries were collected from online sources, PT patients, and experts, and were then submitted to ChatGPT-4. Three board-certified experts independently assessed the accuracy and comprehensiveness of the responses. Readability was measured using the Flesch-Kincaid Grade Level (FKGL: higher scores indicate a higher grade reading level). The Patient Education Materials Assessment Tool (PEMAT) evaluated understandability and actionability (0-100%, higher scores indicate information with clearer messages and more identifiable actions). Semantic Textual Similarity (STS score, 0-1; higher scores indicate higher similarity) assessed variation in the meaning of texts over two months (including ChatGPT-4o) and for different terminologies related to PT.

Results: Sixteen (33%) of the 48 responses were rated accurate, while 36 (75%) were rated comprehensive. Only 17% of treatment-related questions received accurate responses. Most responses were written at a college reading level (median and interquartile range [IQR] of FKGL score: 15.4 [14.4-16.6]). The median of PEMAT for understandability was 83% (IQR: 70%-92%), and for actionability, it was 60% (IQR: 40%-60%). The medians of STS scores in the meaning of texts over two months and across terminologies were all ≥ 0.9.

Conclusions: ChatGPT-4 provided generally comprehensive information in response to patient-centered queries but lacked accuracy and was difficult to read for individuals below a college reading level.

Introduction

ChatGPT, a large language model (LLM) released by OpenAI in November of 2022 (OpenAI, 2022), stands at the forefront of the rapidly evolving field of artificial intelligence (AI). Utilizing a self-supervised learning approach and training on a diverse range of internet texts, it is able to summarize question-targeted information in a human-like and plain language (Anderson et al., 2023; Ayre et al., 2024; Haupt and Marks, 2023; Lyu et al., 2023). In this regard, ChatGPT has garnered enormous attention in medicine (Liu et al., 2023; Homolak, 2023; Will chatgpt and transform healthcare, 2023; Thirunavukarasu et al., 2023), and early research found promising results in its ability to address generally accurate patient queries in areas such as urology (Davis et al., 2023), myopia (Lim et al., 2023), and cancer (Pan et al., 2023). In addition, it has demonstrated a potential performance on successfully completing orthopedic examinations (Kung et al., 2023; Fiedler et al., 2024), and development of a personalized exercise prescription (Cavazzotto et al., 2023). With the increasing reliance on online sources for medical information (Kuehn, 2013), people are likely to turn to this AI tool for similar queries (Will chatgpt and transform healthcare, 2023).

Patients with patellar tendinopathy (PT) often experience longstanding pain (Scott et al., 2020) and consider a diverse range of treatment options (Millar et al., 2021). Despite communicating with healthcare providers, patients can still have lingering questions afterward about their initial diagnosis, treatment options, and time to recover due to ambiguous underlying pathology and treatment effectiveness (Millar et al., 2021; Rosen et al., 2022). Situations requiring instant advice are also common, such as when prescribed exercises cause a flare-up of pain (Millar et al., 2021; Sprague et al., 2021). Determining whether ChatGPT is a valuable tool for PT patients to seek information is imperative. If this easy-to-access free tool could provide timely and appropriate advice, it may improve health literacy and self-management, as well as bridge the gap between patients and clinicians. However, there are currently no studies exploring the application of this AI tool for patients with PT.

Therefore, we aimed to evaluate the performance of ChatGPT-4 in responding to PT patient-centered questions by (1) assessing accuracy, comprehensiveness, readability, understandability, and actionability; (2) comparing these metrics for questions within five domains, namely epidemiology/pathogenesis/etiology, clinical presentation, diagnosis, treatment, and prognosis; and (3) exploring the variability of its output over time and between different terminologies used for PT.

Methods

Design

This cross-sectional study was conducted between September 25, 2023, and May 2, 2024, at the Erasmus MC University Medical Center. Fig.1 illustrates the study design. The protocol for this study was pre-registered on Open Science Framework (https://doi.org/10.17605/OSF.IO/Z2W7P). Local ethical approval was obtained for this study (MEC-2016-500), and all patients provided written informed consent.

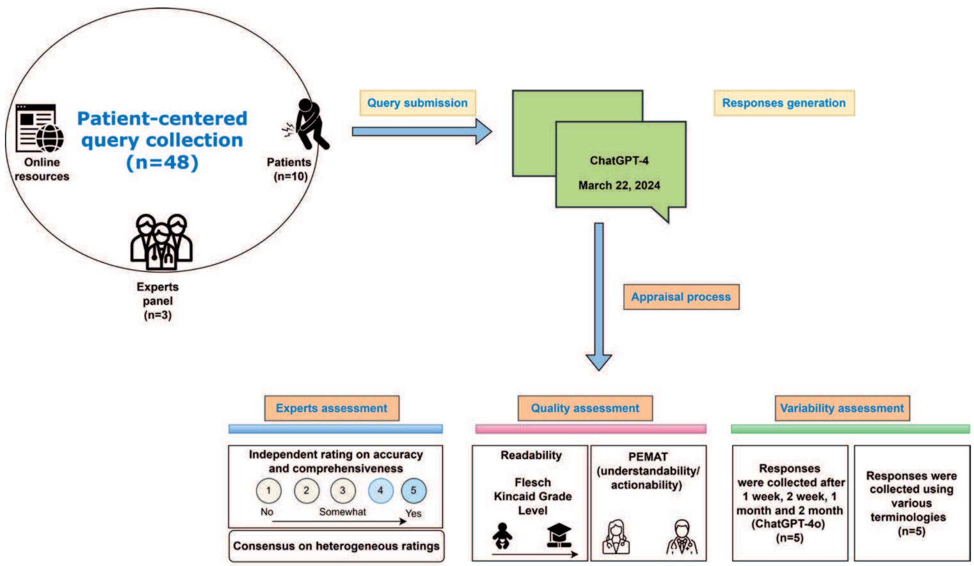


Figure 1. Flow chart of the study design

Icons were created with image: Flaticon.com

Participants

Before consulting PT patients, a list of commonly asked questions for PT was collected from online resources. Ten athletes (median age: 33 years; 30% female) with clinically diagnosed PT confirmed by ultrasound reviewed and edited this query. Patients from the JUMPER study (Breda et al., 2021), a randomized clinical trial (RCT) comparing 2 exercise therapies for PT, who were scheduled for follow-up visits during this study period were asked to participate in the current study. Inclusion criteria were: age 18–35 years; performing sports ≥ 3 times a week; patellar tendon pain increasing with activity; tenderness on palpation; structural tendon changes on ultrasound, and/or increased vascularity on Doppler. An expert panel (n = 3) was selected based on their extensive publication records, identified through the Web of Science using the search terms 'tendinopathy' and 'PT.' The experts (R-JdV, KS and PM) refined the queries and independently assessed the accuracy and comprehensiveness of the ChatGPT responses.

Protocol deviation

Due to the launch of ChatGPT-4 omit (GPT-4o)(OpenAI. Hello gpt, 2024) on May 13, 2024, which is purportedly optimized for real-time responses, we supplemented our variability assessment with responses generated on May 22, 2024. These responses, gathered two months after the initial submission, were not included in the published protocol. We decided to add this analysis to improve the comprehension of the impact of this rapid development within the study period.

Patient-centered query collection

The process began with online searching using the search term 'patellar tendinopathy' on Google Trends over the past five years (Table S1A). Then, we sourced queries from reputable health information outlets or websites using the 'site:' operator in Google, including the National Health Service (patellar tendinopathy site:nhs.uk), the National Institute of Arthritis and Musculoskeletal and skin Diseases (patellar tendinopathy site:niams.nih.gov) and Mayo Clinic (patellar tendinopathy site:mayoclinic.org). Queries on the outlet or website focusing on PT or synonyms were reviewed and extracted. Repeated queries were excluded. Additionally, questions about PT posted on an online health-related forum (Reddit's/AskDocs)(Ask docs) by September 25, 2023, were collected. Collectively, these queries were then categorized into five pre-specified domains: epidemiology/pathogenesis/etiology, clinical presentation, diagnosis, treatment, and prognosis. Next, 10 PT patients reviewed and edited this preliminary query to ensure it covered the topics of interest. The expert panel refined the queries based on their clinical experience. Details are in (https://doi-org.eur.idm.oclc.org/10.17605/OSF.IO/Z2W7P). Finally, 48 English questions were selected (Table S2A). The whole process of query collection was English-based. It excluded questions in other languages and from other geographical regions.

Generation output from ChatGPT-4

Each 'standalone' question was submitted to ChatGPT-4 on March 22, 2024, and the corresponding responses are in Table S3A.

Outcomes measurements

Accuracy and comprehensiveness

From April 4 to April 20, 2024, the expert panel independently rated the accuracy and comprehensiveness of each response based on their experiences due to lack of an internationally published guideline or consensus statement on PT management. We used a 5-point Likert scale (Davis et al., 2023; Lim et al., 2023) as follows for accuracy 1) ‘Very inaccurate/Misleading’ for responses entirely being incorrect or distorting; 2) ‘Inaccurate’ for responses with significant errors; 3) ‘Somewhat accurate’ for responses being partially correct; 4) ‘Accurate’ for responses essentially being correct and factual; 5) ‘Very accurate’ for responses being entirely accurate and precise. For comprehensives, the scale was used as follows: 1) ‘Very incomprehensive’ for response extremely lacking details; 2) ‘Incomprehensive’ for responses lacking essential details; 3) ‘Somewhat incomprehensive’ for responses covering basic but incomplete details; 4) ‘Comprehensive’ for responses covering most necessary details; 5) ‘Very comprehensive’ for responses providing exhaustive details. The average rating across three experts was registered (1–5 points). Responses rated as 4 or 5 points by all experts were graded as ‘accurate’ or ‘comprehensive’, while those rated as 1–3 by all experts were classified as ‘inaccurate’ or ‘incomprehensive,’ with explanations provided. For responses with mixed ratings, an online consensus meeting was held where experts discussed and reached an agreement on their classification. The results, including all inaccurate and incomprehensive ratings (1–3 points by all experts), and their corresponding explanations, were documented in Table S4A.

Quality assessment

Readability was evaluated using the online calculator (https://www.readabilityformulas.com), which was embedded in formula Flesch-Kincaid Grade Level (Kincaid et al., 1975; Badarudeen and Sabharwal, 2010) (FKGL), dependent on average words per sentence and syllables per word. Higher FKGL scores indicate higher reading ability needed to understand the text (e.g., a score of 12.5 [a 12th grade or first year of college], given as a U.S. school grade level). Because relying on the readability formula neglects many of the factors that contribute to comprehension and can be misleading (Lipari et al., 2019; Agency for Healthcare Research and Quality, 2020a), Two researchers (JD and JJ) used the Patient Education Materials Assessment Tool (PEMAT) (Agency for Healthcare Research and Quality, 2020a) for printable material (PEMAT-P) (Agency for Healthcare Research and Quality, 2020b) to independently evaluate the understandability (17 items) and actionability (7 items) of ChatGPT-4 responses. In case of disagreement, a consensus was achieved through a discussion between the reviewers. If disagreements persisted, the evaluation would be resolved by a third researcher (R-JdV). Understandability refers to how easy it is for patients to grasp key messages, while actionability is defined as how well the information guides patients in identifying actions they can take based on the information provided (Agency for Healthcare Research and Quality, 2020a). Scores range from 0% to 100%, with higher scores indicating greater clarity and guidance. Inter-rater reliability (0.92 and 0.93, respectively) was excellent in a previous study (Vishnevetsky et al., 2018). Detailed descriptions are in Table S5A.

Variability assessment

Given the inherent stochastic and free-text generation feature of LLMs, we tested whether ChatGPT-4 responses differed over two months and across terminologies (‘patellar tendinopathy’, ‘patellar tendinitis’, ‘patellar tendinosis’, ‘Jumper’s knee’). Five questions from the query set were analyzed (https://doi-org.eur.idm.oclc.org/10.17605/OSF.IO/Z2W7P). Variations were assessed using the Semantic Textual Similarity (STS) framework (Agirre et al., 2013), measuring how closely two texts align in meaning beyond their literal expressions. STS scores, ranging from 0 to 1, were computed using a pre-trained model of Sentence Transformers (SBERT)(Reimers, 2019), a state-of-the-art deep learning model for natural language processing (NLP) tasks. Scores closer to 1 indicate greater similarity in meaning between two texts (Eq. (A.1).

Statistical analysis

Descriptive statistics were used to summarize the accuracy, comprehensiveness, readability, understandability, actionability, and STS scores. Data distribution was tested using the Shapiro-Wilk test and the histogram graph. Mean with standard deviation (SD) was reported for normally distributed data; otherwise, median with interquartile range (IQR) was used. Frequencies and percentages were displayed for categorical data. For comparison across five domains, the Fisher exact test compared proportions of accurate and comprehensive responses; the Kruskal-Wallis test was used for readability, understandability, and actionability. A 2-tailed test with a p value < 0.05 was considered statistically significant. Inter-rater reliability (IRR) for accuracy and comprehensiveness was calculated using Fleiss’ Kappa; percent agreement was used for PEMAT. Analyses were performed in R studio version 2023.12.0 (RStudio IDE, Boston, MA, USA) with ‘irr’ and ‘ggplot2’ packages. Details of STS score computation are in Eq. (A.1).

Results

Word count of responses

A total of 48 patient-centered questions and responses were collected (Table S2A). The mean word count of the responses was 290 (SD: 70), and the average number of words per sentence was 22 (3) (Table 1).

Table 1. The Summary of assessment in text-length, readability, understandability and actionability

Measure	N	Mean (SD) / Median [IQR]	Range (min-max)
Response total words	48	290 (70)	(161-429)
Response total sentences	48	13 [11-16]	(5-27)
Number of words per sentence	48	22 (3)	(14-32)
Number of syllables per word	48	1.9 [1.8-2.0]	(0-2.1)
FKGL[a]	48	15.4 [14.4-16.6]	(6.4-21.0)
PEMAT understandability score, %[b]	48	83 [70-92]	(58-92)
PEMAT actionability score, %[b]	48	60 [40-60]	(0-60)

Abbreviation: FKGL = Flesch-Kincaid Grade Level; IQR = Interquartile range; min = minimal value; max = maximal value; N = numbers of questions; PEMAT = The Patient Education Materials Assessment Tool; SD = standard deviation.

[a]Scored from 6 (6th grader to reading level) to 21 (post-grad level to reading level).
[b]Scored from 0% (low) to 100% (high).

Expert assessment

Figure. 2A displays the overview of the distribution of ratings across three experts. The overall mean (SD) score was 3.29 (0.68) for accuracy and 3.78 (0.56) for comprehensiveness (Figure. 2B). Of the 48 responses, 7 (15%) and 13 (27%) of them were rated as accurate and comprehensive, respectively, receiving scores of 4 or 5 points by all three experts (IRR of 0.16 [p = 0.057] for accuracy and 0.06 [p = 0.48] for comprehensiveness). After reaching a consensus, the proportion of responses rated as comprehensive increased to 75% (n=36), while those rated as accurate increased to 33% (n=16), as shown in Figure. 2C. There was a significant difference in the proportions of accurate responses across the five domains (p = 0.032). Notably, only 4 (17.4%) of the responses to treatment-related questions were rated as accurate (Table 2). No statistically significant difference was observed in the proportions of comprehensive responses across five domains (p=0.094).

Quality assessment

Table 1 shows the median of the FKGL score for readability was 15.4 (IQR: 4.6-16.6), indicating extremely difficult reading material that requires at least a college level to understand. No statistically significant difference was found in FKGL scores among the five categories (p = 0.215). The median of the PEMAT understandability score was (83% [IQR: 70%-92%]; IRR: 93%), implying that ChatGPT's responses were likely to be easily comprehended by patients. The median of the PEMAT actionability score was 60.0% (IQR: 40.0%-60.0%); IRR: 92%), indicating that moderately actionable information that patients could identify. When comparing each score across five domains

(Table 2), there was a significant difference in understandability (p = 0.023), though not in actionability (p = 0.090).

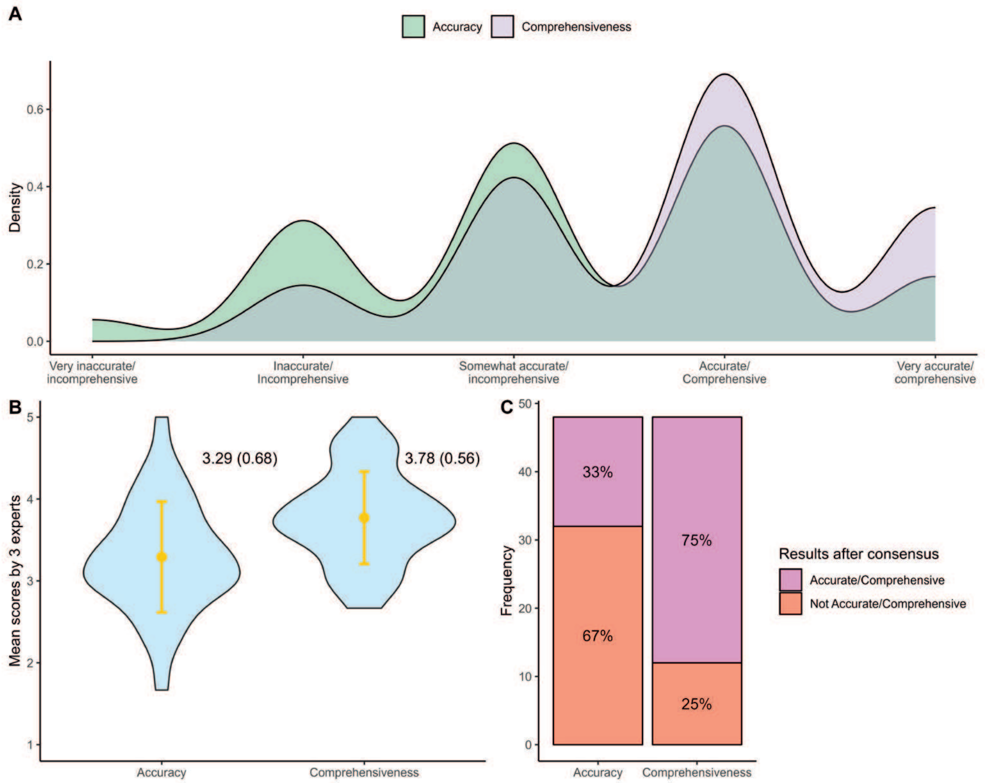


Figure 2. Overall assessment of responses rated as accurate and comprehensive by 3 experts

A: the distribution of ratings (a 5-point Likert scale) by three experts using a Kernel density plot.
B: the average mean scores by three experts using violin plots, which can visualize the distribution of means. Dots represent mean value, bars represent standard deviations. Panel C: the proportion of accurate or comprehensive responses after achieving consensus.

Table 2. Sub-analysis in comparisons among 5 domains

Domain (n=48)	After consensus; n(%)		FKGL; median [IQR]	PEMAT (0-100%); median [IQR]	
	Accurate[a] (n=16)	Comprehensive (n=36)		Understandability[b]	Actionability
Epidemiology/ Pathogenesis / Etiology (n=6)	2 (33)	5 (83)	15.5 [14.4-16.0]	83.0 [68.8-83.0]	40 [5-60]
Clinical presentation (n=4)	2 (50)	4 (100)	14.6 [13.9-14.8]	67.0 [64.8-69.0]	40 [20-55]
Diagnosis (n=6)	5 (83)	6 (100)	16.6 [16.4-17.9]	81.5 [76.3-83.0]	20 [0-55]
Treatment (n=23)	4 (17)	13 (57)	15.9 [14.2-17.4]	83.0 [72.5-85.0]	60 [60-60]
Prognosis (n=9)	3 (33)	8 (89)	15.2 [15.0-16.0]	92.0 [92.0-92.0]	60 [60-60]

Abbreviation: FKGL = Flesch-Kincaid Grade Level; IQR = Interquartile range. PEMAT = The Patient Education Materials Assessment Tool.

[a]p = 0.032, [b]p = 0.023, comparison among 5 domains.

Variability assessment

The medians of the STS scores for responses at the subsequent four-time points were all above 0.9 (Figure. 3A), suggesting that the texts generated within two months had substantially similar meanings (Table S6A). Additionally, high STS scores (median range: 0.90-0.93) were observed in the meaning of responses to questions using other terminologies (Figure. 3B). This implied ChatGPT-4 tends to be adaptive to different terminologies of PT (Table S7A).

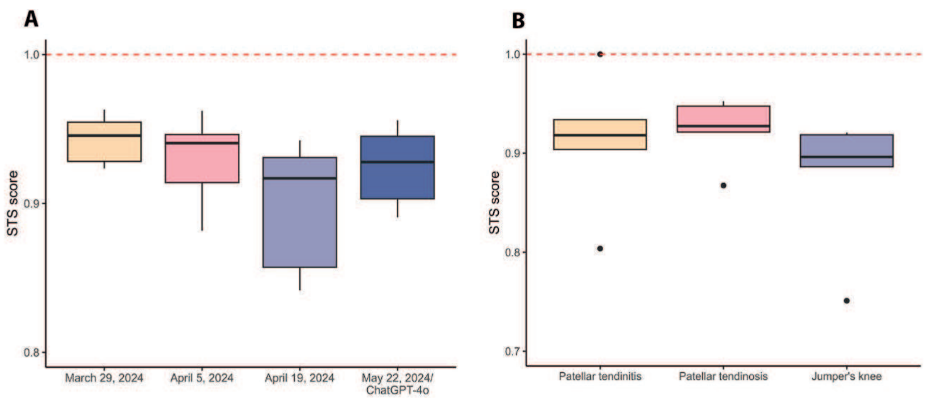


Figure 3. Semantic textual similarity (STS) scores on questions at different time points and different terminologies.

The summary of STS scores is displaying using boxplots. Lines in the box represent median, the box margins represent upper quartile and lower quartile respectively. Dots outside the box represent outliers. A: Red dash lines represent the reference: responses from initial submission on March 22, 2024. B: responses given the question using 'patellar tendinopathy'.

Discussion

This is the first study that benchmarked the performance of ChatGPT-4 in answering patient-centered queries for tendinopathy. We found that ChatGPT-4 can provide considerably comprehensive, understandable, and actionable responses to patients' questions related to PT, whereas the accuracy and readability were notably low. Specifically, ChatGPT-4 performed poorly in delivering accurate treatment-focused questions. We also shed light on the consistent performance of ChatGPT-4 over time, and adaptive ability to different terminologies used for PT. As generative AI or LLMs are entering the mainstream of clinical medicine (Will chatgpt and transform healthcare, 2023; Thirunavukarasu et al., 2023), healthcare providers should stay informed about the capabilities and limitations of AI tools like ChatGPT. Its strengths and shortcomings must be clarified and addressed for specific patient populations.

Accuracy and readability

The application of ChatGPT-4 or other LLMs in tendinopathies has been rarely described. To our knowledge, only one study has evaluated this tool in patients with rotator cuff tears, however, the authors did not involve patients in the study design and refrained from assessing the accuracy of the generated answers (Günay et al., 2024). ChatGPT's ability to answer queries related to diseases such as myopic (Lim et al., 2023), thyroid nodules (Campbell et al., 2024) and anterior cruciate ligament (ACL) surgery (Kaarre et al., 2023) has been partially explored. In contrast to our findings on accuracy (33%), approximately 65%–80% of ChatGPT's responses were rated accurate in the aforementioned studies. The reason for this may be that our study collected a larger number of queries which may reveal more variability in its performance. More importantly, the rating of these responses can be subjective. In this current study, we found a considerable dispersion (Fig. 2B) and a low IRR among experts rating when using a 5-point Likert scale, emphasizing the need for caution in interpreting these findings. Due to the absence of an internationally published clinical guideline for PT, experts had to rate the accuracy of these responses based on existing literature and their own experiences by comparing the outputs generated by LLMs to human approaches. Moreover, we took a stringent consensus approach to rate accuracy, even rating responses as inaccurate if ChatGPT's answers appeared overly optimistic for patients despite being generally accurate (Table S4A). This strict criteria may explain why we reported a lower accuracy rate than the aforementioned studies.

Our results demonstrated that ChatGPT-4 performed extremely poorly in delivering accurate treatment-targeted responses. For example, when prompted with "what is the best treatment in PT," ChatGPT-4 generated a list of treatments without prioritizing or ranking them. Additionally, while our understanding of the optimal first-line treatment for PT is progressive tendon-loading exercises (Breda et al., 2021; Marigi et al., 2022), this tool frequently provided general, non-individualized options, with responses often starting with rest, ice, and eccentric exercises. It even raises concerns that this AI tool is prone to recommending certain injection treatments, such as platelet-rich plasma, despite lack of evidence (Challoumas et al., 2021; Scott et al., 2019). This tendency may be attributed to the presence of advocating such treatments by third parties, such as online advertisements or non-evidence-based sources, included in the data on which ChatGPT was trained (OpenAI, 2022). These findings underscore the risk of PT patients receiving misinformation from ChatGPT-4, particularly in the area of treatment recommendations.

For readability, we found ChatGPT-4 generating texts were written at the college reading level based on FKGL, exceeding the recommended sixth grade level (Weiss, 2003) for patient education materials. Consistent with other studies (Davis et al., 2023; Pan et al., 2023; Günay et al., 2024; Cocci et al., 2024), these findings indicate this AI tool used medical information that may not be useful for lay audience. However, it is important to note that this limitation might be partially mitigated by customizing prompts to request a lower reading level (Ayre et al., 2024; Campbell et al., 2024; Rouhi et al., 2024), although most of these responses still fail to meet the targeted readability, especially in the old version of ChatGPT. Collectively, these findings highlight future efforts to reduce the complexity of LLMs-generated text, which may be crucial for improving patient education materials. This also suggests that patients may need additional support or guidance to formulate effective prompts.

Comprehensiveness, understandability and actionability

ChatGPT-4 demonstrated an overall good ability in addressing comprehensive content about PT. Several studies have also reported this ability in delivering comprehensive content for care of other diseases (Davis et al., 2023; Lim et al., 2023). These findings would be not so surprising as this so called 'all-encompassing' model had been trained on billions of words derived from scientific articles, books and other internet-based content (OpenAI, 2022; Thirunavukarasu et al., 2023). However, it should be noted that despite sounding authoritative, it lacks uncertainty weightings on recommendations and references.

Additionally, our results exhibited that ChatGPT-4 showed good performance in delivering generally understandable and actionable information for patients to grasp key messages and identify what they can do. A similar high performance in these two domains was reported in a study evaluating the quality of ChatGPT-4 in plastic surgery consultations (Yun et al., 2023). However, compared to their results, we observed a relatively lower score in actionability (93% vs. 60%). This discrepancy seems reasonable, given that the calculation of actionability score is highly dependent on whether the query sought recommendations. Compared with only 48% (23/48) of our queries focused on treatment advice, their queries were primarily related to decision-making.

Clinical implications

Healthcare providers currently should not guide or advise PT patients to rely onthese tools independently, especially for treatment advice. People often cannot distinguish between information provided by LLMs and human clinicians (Thirunavukarasu et al., 2023), and tend to trust responses that mimic human behaviour (Will chatgpt and transform healthcare, 2023). If patients choose to use such tools, healthcare providers should proactively inform them about these potential risks, such as misinformation or treatment recommendations with limited evidence or even evidence against their use.

However, ChatGPT-4 showed potential ability to generate understandable, actionable, and comprehensive PT content. Policy makers and healthcare professionals could in the future leverage this tool to supplement clinical task such as PT patient education, particularly in areas where clear and concise information is needed. More insight into the sources used to deliver information might help to further improve the accuracy of the responses and optimizing prompts for questions related to healthcare for PT might be evaluated to enhance the readability of the responses. Future studies are warranted in evaluating if patients can use this tool as a starting point for discussion before they consult healthcare providers. Nevertheless, Healthcare providers must verify the accuracy and readability of the AI-generated content to ensure it meets the needs of their patients.

Strengths and limitations

This study has several strengths as we employed a robust patient-centered study design, including a multi-source collection approach, strict reviews by three experts, and multilevel assessments. Additionally, our study utilized innovative deep learning-based similarity analysis to allow us to investigate both temporal consistency and terminological flexibility of ChatGPT. Nonetheless, this study has certain limitations that must be acknowledged. Currently, there are no validated tools specifically designed to assess the quality of LLMs outputs. While PEMAT is a validated tool to assess patient education materials, it may be inappropriate or insufficient to fully capture the unique aspects of quality specific to AI-generated responses. We only determined the ChatGPT's ability in addressing authoritative PT questions. However, we did not explore whether these outputs can be refined by an iterative process, such as adding instructions to prompt for references, lower readability, or even peer-reviewed sources. Furthermore, we did not compare ChatGPT's performance with other LLMs. The gold standard for accuracy and comprehensiveness was the view of three different experts, which may be challenged as their personal views may be diverse and biased as well, especially the absence of international guidelines for PT. Furthermore, surgeons—who might offer valuable insights, particularly regarding surgical queries—were not included in our expert panel. A more diverse pool of raters, encompassing additional clinical specialties, may further strengthen future evaluations. Additionally, the inclusion of patients in this phase of rating the Chat-GPT responses would have added another perspective. We chose to focus more on quantitative outcomes rather than qualitative data.

Conclusion

While ChatGPT-4 provided generally comprehensive, understandable, and actionable responses to patient-centered queries for PT, it currently performed poorly in delivering accurate and sufficiently readable information. Our findings suggest that the current version of ChatGPT is not accurate enough for PT patients to use independently to seek information. Information from healthcare professionals presently remains the standard of trustable delivery of patient education for PT patients. As generative AI is fast evolving, future research should focus on whether healthcare providers could involve the development of medical LLMs to improve accuracy, and use ChatGPT or other medical LLMs to actually reduce clinical workload of health care professionals.

References

Agency for Healthcare Research and Quality, 2020a. The Patient Education Materials Assessment Tool (Pemat) and User's Guide. Introduction. https://www.ahrq.gov/he alth-literacy/patient-education/pemat1.html. (Accessed 24 March 2024).

Agency for Healthcare Research and Quality, 2020b. The Patient Education Materials Assessment Tool (Pemat) and User's Guide. How to Use the Pemat to Assess a Material. Available from: https://www.ahrq.gov/health-literacy/patient-education/ pemat2.html. (Accessed 24 March 2024).

Agirre, E., Cer, D., Diab, M., Gonzalez-Agirre, A., Guo, W., 2013. *sem 2013 shared task: Semantic textual similarity. Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity. Association for Computational Linguistics, pp. 32–43.

Anderson, L.B., Kanneganti, D., Bentley Houk, M., Holm, R.H., Smith, T., 2023. Generative Ai as a Tool for Environmental Health Research Translation. https://doi.org/10.1101/2023.02.14.23285938 medRxiv.

Ask docs. Reddit. Available from: https://www.reddit.com/r/AskDocs/. (Accessed 25 September 2023).

Ayre, J., Mac, O., McCaffery, K., et al., 2024. New frontiers in health literacy: using chatgpt to simplify health information for people in the community. J Gen Intern Med 39, 573–577.https://doi.org/10.1007/s11606-023-08469-w.

Badarudeen, S., Sabharwal, S., 2010. Assessing readability of patient education materials: current role in orthopaedics. Clin Orthop Relat Res 468, 2572–2580. https://doi.org/10.1007/s11999-010-1380-y.

Breda, S.J., Oei, E.H.G., Zwerver, J., et al., 2021. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. Br J Sports Med 55, 501–509. https://doi.org/10.1136/bjsports-2020- 103403.

Campbell, D.J., Estephan, L.E., Sina, E.M., et al., 2024. Evaluating chatgpt responses on thyroid nodules for patient education. Thyroid 34, 371–377. https://doi.org/ 10.1089/thy.2023.0491.

Cavazzotto, T.G., Dantas, D.B., Queiroga, M.R., 2023. Chatgpt and exercise prescription: human vs. Machine or human plus machine? J Sport Health Sci 13, 661–662. https://doi.org/10.1016/j.jshs.2023.10.008.

Challoumas, D., Pedret, C., Biddle, M., et al., 2021. Management of patellar tendinopathy: a systematic review and network meta-analysis of randomised studies. BMJ Open Sport Exerc Med 7, e001110. https://doi.org/10.1136/bmjsem-2021- 001110.

Cocci, A., Pezzoli, M., Lo, Re M., et al., 2024. Quality of information and appropriateness of chatgpt outputs for urology patients. Prostate Cancer Prostatic Dis 27, 103–108. https://doi.org/10.1038/s41391-023-00705-y.

Davis, R., Eppler, M., Ayo-Ajibola, O., et al., 2023. Evaluating the effectiveness of artificial intelligence-powered large language models application in disseminating appropriate and readable health information in urology. J Urol 210, 688–694. https://doi.org/10.1097/ju.0000000000003615.

Fiedler, B., Azua, E.N., Phillips, T., Ahmed, A.S., 2024. Chatgpt performance on the american shoulder and elbow surgeons maintenance of certification exam. J Shoulder Elbow Surg 33, 1888–1893. https://doi.org/10.1016/j.jse.2024.02.029.

Günay, A.E., O¨zer, A., Yazıcı, A., Sayer, G., 2024. Comparison of chatgpt versions in informing patients with rotator cuff injuries. JSES Int 8, 1016–1018. https://doi.org/ 10.1016/j.jseint.2024.04.016.

Haupt, C.E., Marks, M., 2023. Ai-generated medical advice-gpt and beyond. JAMA 329, 1349–1350. https://doi.org/10.1001/jama.2023.5321.

Homolak, J., 2023. Opportunities and risks of chatgpt in medicine, science, and academic publishing: a modern promethean dilemma. Croat Med J 64, 1–3.

Kaarre, J., Feldt, R., Keeling, L.E., et al., 2023. Exploring the potential of chatgpt as a supplementary tool for providing orthopaedic information. Knee Surg Sports Traumatol Arthrosc 31, 5190–5198. https://doi.org/10.1007/s00167-023-07529-2.

Kincaid Jr., J.P., Rpf, Rogers, R.L., Chissom, B.S., 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Naval Technical Training Command Millington TN Research Branch.

Kuehn, B.M., 2013. More than one-third of us individuals use the internet to self- diagnose. JAMA 309, 756–757. https://doi.org/10.1001/jama.2013.629.

Kung, J.E., Marshall, C., Gauthier, C., Gonzalez, T.A., Jackson, J.B., 2023. Evaluating chatgpt performance on the orthopaedic in-training examination. JB JS Open Access 8. https://doi.org/10.2106/jbjs.Oa.23.00056.

Lim, Z.W., Pushpanathan, K., Yew, S.M.E., et al., 2023. Benchmarking large language models' performances for myopia care: a comparative analysis of chatgpt-3.5, chatgpt-4.0, and google bard. EBioMedicine 95, 104770. https://doi.org/10.1016/j. ebiom.2023.104770.

Lipari, M., Berlie, H., Saleh, Y., Hang, P., Moser, L., 2019. Understandability, actionability, and readability of online patient education materials about diabetes mellitus. Am J Health Syst Pharm 76, 182–186. https://doi.org/10.1093/ajhp/ zxy021.

Liu, J., Wang, C., Liu, S., 2023. Utility of chatgpt in clinical practice. J Med Internet Res 25, e48568. https://doi.org/10.2196/48568.

Lyu, Q., Tan, J., Zapadka, M.E., et al., 2023. Translating radiology reports into plain language using chatgpt and gpt-4 with prompt learning: results, limitations, and potential. Visual Computing For Industry, Biomedicine, and Art 6, 9. https://doi. org/10.1186/s42492-023-00136-5.

Marigi, E.M., Buckley, P., Razi, F., et al., 2022. Patellar tendinopathy: critical analysis review of current nonoperative treatments. JBJS Rev 10 e21.00168.

Millar, N.L., Silbernagel, K.G., Thorborg, K., et al., 2021. Tendinopathy. Nat Rev Dis Primers 7, 1. https://doi.org/10.1038/s41572-020-00234-1.

OpenAI, 2022. Introducing chatgpt. Available from: https://openai.com/blog/chatgpt. (Accessed 22 March 2024).

OpenAI. Hello gpt-4o, 2024. Available from: https://openai.com/index/hello-gpt-4o/. (Accessed 22 May 2024).

Pan, A., Musheyev, D., Bockelman, D., Loeb, S., Kabarriti, A.E., 2023. Assessment of artificial intelligence chatbot responses to top searched queries about cancer. JAMA Oncol 9, 1437–1440. https://doi.org/10.1001/jamaoncol.2023.2947.

Reimers, N.G., 2019. Iryna. Sentence-bert: sentence embeddings using siamese bert- networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics, pp. 3982–3992.

Rosen, A.B., Wellsandt, E., Nicola, M., Tao, M.A., 2022. Clinical management of patellar tendinopathy. J Athl Train 57, 621–631.

Rouhi, A.D., Ghanem, Y.K., Yolchieva, L., et al., 2024. Can artificial intelligence improve the readability of patient education materials on aortic stenosis? A pilot study. Cardiol Ther 13, 137–147. https://doi.org/10.1007/s40119-023-00347-0.

Scott, A., LaPrade, R.F., Harmon, K.G., et al., 2019. Platelet-rich plasma for patellar tendinopathy: a randomized controlled trial of leukocyte-rich prp or leukocyte-poor prp versus saline. Am J Sports Med 47, 1654–1661. https://doi.org/10.1177/ 0363546519837954.

Scott, A., Squier, K., Alfredson, H., et al., 2020. Icon 2019: international scientific tendinopathy symposium consensus: clinical terminology. Br J Sports Med 54, 260–262. https://doi.org/10.1136/bjsports-2019-100885.

Sprague, A.L., Coupp´e, C., Pohlig, R.T., Snyder-Mackler, L., Silbernagel, K.G., 2021. Pain- guided activity modification during treatment for patellar tendinopathy: a feasibility and pilot randomized clinical trial. Pilot Feasibility Stud 7, 58. https://doi.org/ 10.1186/s40814-021-00792-5.

Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K., Gutierrez, L., Tan, T.F., Ting, D.S.W., 2023. Large language models in medicine. Nat Med 29, 1930–1940. https://doi.org/ 10.1038/s41591-023-02448-8.

Vishnevetsky, J., Walters, C.B., Tan, K.S., 2018. Interrater reliability of the patient education materials assessment tool (pemat). Patient Educ Couns 101, 490–496. https://doi.org/10.1016/j.pec.2017.09.003.

Weiss, B., 2003. Health Literacy: A Manual for Clinicians. American Medical Association. Will chatgpt transform healthcare, 2023. Nat Med 29, 505–506. https://doi.org/10.1038/s41591-023-02289-5.

Yun, J.Y., Kim, D.J., Lee, N., Kim, E.K., 2023. A comprehensive evaluation of chatgpt consultation quality for augmentation mammoplasty: a comparative analysis between plastic surgeons and laypersons. Int J Med Inform 179, 105219. https://doi. org/10.1016/j.ijmedinf.2023.105219.

Supplementary material

Supplementary material for this paper can be found online:
https://doi.org/10.1016/j.msksp.2025.103275

3

Chapter 3

Long-term Prognosis of Athletes With Patellar Tendinopathy Receiving Physical Therapy: Patient-Reported Outcomes at 5-Year Follow-up

Jie Deng, Jelle J Oosterhof, Denise Eygendaal, Stephan J Breda, Edwin HG Oei, Robert-Jan de Vos

Am J Sports Med. 2025 Jun;53(7):1568-1576.

Abstract

Background: Patellar tendinopathy (PT) is a highly prevalent injury among jumping athletes. The long-term prognosis of athletes with PT following physical therapy is unknown.

Purpose: To assess self-perceived recovery rate and the 5-year change in pain levels, disability, and sports participation and explore the prognostic factors associated with self-perceived recovery.

Study Design: Cohort study

Methods: Athletes with PT who were previously enrolled in a randomized trial and received education, load management advice and exercise therapy instructions at baseline were eligible. An online questionnaire was sent 5 years after inclusion. Self-perceived recovery was assessed by a dichotomized 7-point global rating of change (recovery was defined as "significantly improved" to "completely recovered"). Pain levels during sports (0 to 10 points) and disability assessed by the Victorian Institute of Sport Assessment-Patellar (VISA-P) score were recorded at baseline and 5 years. Sports participation was categorized into return to performance, return to sport, return to participation and quitting sports. Nonparametric tests were performed to compare scores at baseline and 5 years. Logistic regression models were used to identify prognostic factors.

Results: Of 76 eligible participants, 58 (76%) responded (mean [SD] age, 29 [3.8] years; 28% female). At a mean follow-up of 5 years, 76% of participants felt recovered. Pain levels during sports (median from 7 [IQR, 7-8] to 2 [IQR, 1-4] points) and VISA-P score (median from 57 [IQR, 45-66] to 82 [IQR, 74-97] points) significantly improved from baseline to 5 years (all P <.001). Forty-one participants (71%) returned to their desired sports (68% to performance and 32% below pre-injury level), 12 participants (21%) returned to participation in other sports, and 5 (9%) completely ceased sports participation. Participants who felt unrecovered had higher levels of pain and disability and lower return to performance (all $P < .05$). No prognostic factors were identified that were associated with self-perceived recovery.

Conclusion: Athletes with PT following physical therapy can expect a generally acceptable long-term prognosis. However, almost a quarter did not feel recovered and perceived worse patient-reported outcomes. Clinicians treating athletes with PT may use these findings to estimate the average prognosis.

Introduction

Patellar tendinopathy (PT) is a common injury among athletes performing jumping sports[19], with a prevalence of 21% in basketball players and 25% in volleyball players[25]. This condition is characterized by pain at the inferior pole of the patella[19] and impacts work capacity, sports performance[9] and quality of life[31]. Exercise-based strategies have become the first-line treatment, with various therapeutic modalities as complementary options[21,23].

The overall prognosis of PT in athletes is still uncertain. Limited small-scale studies reported a generally poor prognosis, with 53% of male athletes ceasing their sports career after 15 years[16] and 19% of competitive volleyball players retiring[36]. In contrast, another study found a more optimistic outcome, with 57% of male athletes returning to their pre-injury sport at 3-4 years[1]. However, two of these studies were limited to including males[1,16], with one focusing specifically on volleyball players[36], and their reported outcome measures did not align with current guidelines[11,35]. Most of these studies also reflect the average PT course under clinical approaches from more than 10 to 20 years ago. Given the changes in PT management[6,12,23] and the development of more reliable outcome measurement instruments for tendinopathies under relevant core domains[11,35], it is crucial to update and evaluate the prognosis of PT in the context of current standard care, especially with a larger sample size. Such information is essential, as it provides realistic estimates for future outcomes of PT[13], facilitating patient education and decision-making. Additionally, prognostic factors for the long-term course of PT remain largely unknown, which could help identify potential targets for new interventions and guidance in patient-specific treatment algorithms[28].

Therefore, we conducted this prospective study to estimate 5-year patient-reported outcomes in athletes with symptomatic PT following education, load management advice, and exercise-based approaches. The primary aim of this study was to assess self-perceived recovery rate, change in pain levels, disability and quality of life, and sports participation. A secondary aim was to explore the prognostic factors of self-perceived recovery.

Methods

Study Design and Participants

This prospective cohort study is the 5-year follow-up of a previous randomized controlled trial (RCT)[5] that compared the effectiveness between progressive tendon-loading exercises and eccentric exercises in athletes with PT (detailed exercise protocols were published in the previous work[5]). This trial was prospectively registered for the short-term follow-up (6 months) on ClinicalTrial gov (ID: NCT02938143). This extension observational study was approved by the ethics committee of the Erasmus MC University Medical Center (MEC-2016-500). Since the follow-up was long-term, we expected no ongoing treatment effect of the different exercise programs implemented as part of the original trial. Also, we expected that other treatments might have been provided to some participants. For these reasons we considered it reasonable to combine the two groups into 1 cohort for this longer-term analysis. All participants who volunteered for the 5-year follow-up provided digital informed consent before participation. We adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[37] for reporting observational studies and to the International Consensus Statement for Tendinopathy (ICON)[30] minimum reporting standards for tendinopathy.

The original trial enrolled 76 athletes at Erasmus MC University Medical Center between 2017 and 2019, aged 18 to 35 years. Participants performed their desired sports at least 3 times per week before the injury and were diagnosed with symptomatic PT confirmed by the following criteria: patellar tendon pain increasing with activity; recognizable tenderness on patellar tendon palpation; patellar tendon changes on ultrasound and/or increased vascularity on Power Doppler ultrasound. Demographic data (e.g., sex, age, body mass index [BMI] and symptom duration) and several physical tests (e.g., quadriceps strength) were collected at baseline[5,10], and current height and weight were also collected at 5-year follow-up. For this follow-up study, we attempted to consecutively contact all athletes by email or phone from 2022 to 2024, which was 5 years (after their start of enrollment). For those who agreed to participate and provided digital informed consent, the questionnaire was distributed via a link by email.

Outcome Measures

All participants completed an online questionnaire with questions that can be found in the Supplementary file (Appendix A).

Self-perceived recovery

Participants were asked about their self-perceived recovery compared to the start of the trial. This outcome was designed using the global rating of change (GROC) with a 7-point Likert scale according to the most recent guideline in tendinopathy[11,35], ranging from '3 = completely recovered' to '-3 = worse than ever'. The responses were dichotomized, with recovery defined as "significantly improved" to "completely recovered," whereas those rated as "slightly improved" to "worse than ever" were deemed to have not recovered[18,33].

Pain levels, disability, and quality of life

The pain levels were assessed during activities of daily living (ADL), and the most recent sports activity was measured on a scale of 0 to 10 points, with 0 representing no pain and 10 representing severe pain. Complete relief from pain was defined as a pain level equal to 0 at the 5-year follow-up. To assess disability and quality of life, the Victorian Institute of Sport Assessment-Patella (VISA-P) score[38] and health-related quality of life by European Quality of Life-3 Dimensions (EQ-5D-3L) index[20] were used, respectively. The percentage of participants who achieved the known minimal clinically important difference (MCID) for these scores was reported. The MCID for the change of pain levels (greater than 1.2 points[7]) and VISA-P[14] (greater than 13 points) was used for this purpose. All scales were administered in their Dutch versions and were measured at baseline (start of enrollment) and 5-year follow-up.

Sports participation

The pre-injury level of sports activity was recorded at baseline (TABLE 1). Participants were instructed to complete their current sports participation (type, duration, and frequency per week) in the 5-year follow-up digital questionnaire (Appendix A). The physical demands of sports were categorized from high to low intensity as previously defined in the previous literature[4]: jumping, hard pivoting, and cutting (e.g., basketball, volleyball, football), running, twisting, and turning (e.g., racquet sports, baseball, hockey), and no running, twisting, or jumping (e.g., cycling, swimming).

Sports participation was categorized into (1) return to performance (participating the desired sport at or above the pre-injury level), (2) return to sport (participating the desired sport below pre-injury level), (3) return to participation (modified the desired sport to another sport), and (4) completely quit sports participation, according to the most recent consensus[2]. The rate of return to desired sport was defined as the proportion of participants who resumed their pre-injury sports at 5 years, regardless of performance level (combined category 1 and 2). Participants who returned to participation or completely stopped sports participation were subsequently instructed to choose the most crucial reason for this change, including PT symptoms, fear of re-injury of PT, and non-PT related reasons (e.g., change of interest, lack of time).

Health care consumption

Participants were instructed to select which treatments they received between the end of the trial and the 5-year time point using designed checkboxes and open-ended questions (Appendix A).

Proposed prognostic factors

Predefined baseline variables were used to evaluate the prognostic value of self-perceived recovery. All potential prognostic factors were assessed at baseline. These were sex, age, BMI, symptom duration (weeks), VISA-P score, and quadriceps muscle strength (the maximal isometric voluntary contraction measured by a handheld dynamometer in Newton/kg)[5,10].

Statistical Analysis

We performed the analyses based on the complete case, presenting descriptive statistics as numbers and percentages for categorical data (e.g., self-perceived recovery and sports participation) and central tendency and dispersion for continuous data. For normally distributed data inspected using histogram and Shapiro-Wilk test, mean and standard deviation (SD) were used; otherwise, median and interquartile range (IQR) were presented. We decided to perform a subgroup analysis of participants with self-perceived recovery and non-recovery for the patient-reported outcomes (pain levels, VISA-P score, EQ-5D-3L and sports participation). Between-group comparisons were assessed using t-tests for normalized continuous data. For non-normalized data, the Wilcoxon signed-rank test was performed for paired data, and the Mann-Whitney U test was used for non-paired data. The Fisher's exact test was used for categorical comparison. Logistic regression models were used to assess the prognostic value of predefined baseline variables for events of self-perceived unrecovered. These were done using initial univariable analyses, and variables with $p<.15$ were included[27] in the multivariable model adjusted with treatment allocation[24] from the previous RCT[5]. All calculations were performed using R studio version 2023.12.0. P values $< .05$ were considered statistically significant.

Results

Of the 76 athletes included at baseline, 58 (76%) completed the online questionnaire at a mean follow-up of 5 years (SD, 0.2). At follow-up, the mean age was 29.5 years (SD, 3.8), 28% were female, and the mean height and weight were 185 cm (SD, 10) and 84 kg (SD, 14), respectively. The baseline characteristics for these participants at the start of the JUMPER study are displayed in TABLE 1. Of the 18 nonparticipants, 3 (16%) refused to participate in the follow-up assessment, and 15 (79%) did not respond to phone or two email reminders within 4 weeks (Figure 1). There was no significant difference in baseline variables between responders and non-responders (Appendix B, Table B1).

Most participants (95%) received or applied conservative treatment approaches between the end of the trial (6 months follow-up) and the 5-year follow-up. One participant received patellar tendon debridement surgery at 2.5 years (TABLE 2).

Table 1. Baseline Characteristic of Participants (n=58)

Characteristics	Mean (SD) / Median [IQR] / No. (%)
Demographics	
Age, y	24 (3.8)
Male	42 (72%)
Height, cm	184 (9.6)
Weight, kg	82 (12)
BMI, kg/m^2	23.8 (2.6)
Tendinopathy descriptors	
Symptom duration, wk	106 [52-247]
VISA-P score (0-100)	57 [45-66]
VAS during a single-leg decline squat (0-10)	5 [3-7]
Imaging for assisting diagnosis	
Use of ultrasound and MRI	58 (100%)
Physical test results	
Quadriceps muscle strength, N/kg	4.7 (1.0)
Sports characteristics before the injury	
Physical demands of desired sport[a]	
Jumping, hard pivoting, cutting (e.g., basketball, volleyball)	53 (91%)
Running, twisting, turning (e.g., racquet sports)	5 (9%)
No running, twisting, jumping (e.g., cycling, swimming)	0 (0%)
Frequency of sports activity, d/wk	3 [3-3]
Adjustment of desired sport at baseline	
Stopped	21 (36%)
Decreased	22 (38%)
Equal	15 (26%)

Abbreviations: SD: standard deviation; y: years; BMI: body mass index; wk: weeks; IQR: interquartile range; d/wk: days per week; VAS: Visual Analogue Scale; VISA-P: Victorian Institute of Sport Assessment-Patella (VISA-P); N/kg: Newton/kilograms.

[a]the category was based on a previous literature[4].

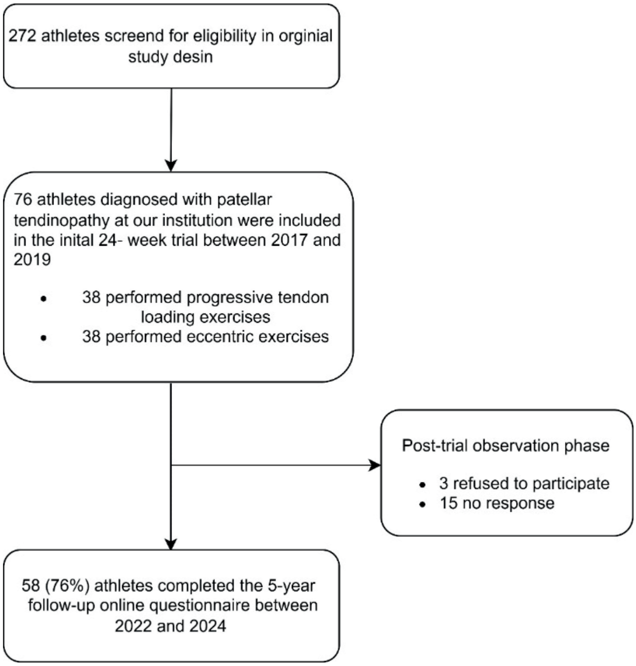


Figure 1. Flow Chart of Study Design

Table 2. Treatment Approaches that Participants Undergone Between the 6-Month Follow-up and the 5-Year Follow-up (58 Participants)

Items	No. (%)
Rest	37 (63)
Adjustment of sports activities	40 (69)
Stretching exercises	25 (43)
Strengthening exercises	42 (72)
Foot orthoses	12 (21)
Knee brace	6 (10)
Patellar strap	15 (26)
Medical taping	9 (16)
Manual therapy	1 (2)
Shock-wave therapy	7 (9)
Ultrasound therapy	4 (7)
Dry needling	6 (10)
Medication (paracetamol and/or anti-inflammatory agents)	8 (14)
Surgery (patellar tendon debridement)	1 (2)

Self-Perceived Recovery

Of the 58 participants, 44 (76%) participants were defined as recovered at 5 years, with 16 (36%) stating complete recovery and 28 of the participants (63%) reporting significant improvement. Fourteen (24%) of the participants did not feel recovered. Among those who were categorized as non-recovered, none reported a worse condition compared to baseline (TABLE 3).

Table 3. Self-Perceived Recovery Using the Global Rating of Change (GROC) with 7-Point Likert Scale[a]

Item	No. (%)
Recovered	44 (76)
Completely recovered	16 (28)
Significantly improved	28 (48)
Non-recovered	14 (24)
Slightly improved	8 (14)
Remain the same	6 (10)
Slightly worsened	0 (0)
Significantly worsened	0 (0)
Worse than ever	0 (0)

[a]Global rating of change was dichotomized into “recovered” and “non-recovered.”

Change in Pain Levels, Disability and Quality of Life

There was a significant improvement in the median scores for pain levels, disability level, and quality of life from baseline to 5-year follow-up (all *P* values < .001). The proportion of participants who achieved the MCID at 5 years was 61% for pain levels during ADL, 67% for pain levels during recent sports activity, and 71% for the VISA-P score. Additionally, at the 5-year follow-up, 21 participants (36%) reported complete relief from pain during ADL, and 13 (25%) were completely pain-free during sports activity (TABLE 4).

In addition, there were significant differences in the median of these change scores between those who felt recovered and non-recovered (Appendix B, Table B2).

Table 4. Comparison of Pain Levels, Disability and Quality of Life Between Values at Baseline and 5 Years

Variables	At baseline, median [IQR]	At 5-year, Median [IQR] / No. (%)	*P* value	MCID achieved (%)
Pain levels during ADL (0-10 points)	4 [3-6]	1 [0-3]	< .001	61
Complete relief from pain (0 points)		21 (36)		
Pain levels during recent sports activity (0-10 points)	7 [7-8]	2 [1-4][a]	< .001	67
Complete relief from pain (0 points)		13 (25)[a]		
VISA-P score (0-100 points)	57 [45-66]	82 [74-97]	< .001	71
EQ-5D-3L index (0-1)	0.84 [0.81-0.84]	0.98 [0.90-0.99]	< .001	NA

[a]A total of 53 participants rated their pain levels due to 5 having quit sports at 5 years.

Abbreviations: IQR: interquartile range; VISA-P: the Victorian Institute of Sport Assessment-Patellar; VAS: Visual Analogue Scale; EQ-5D-3L: European Quality of Life-3 Dimensions; MCID: The minimal clinically important difference; NA: non-applicable.

Sports Participation

The types of participated sports at pre-injury and 5-year follow-up were displayed based on level of physical demand (Figure 2).

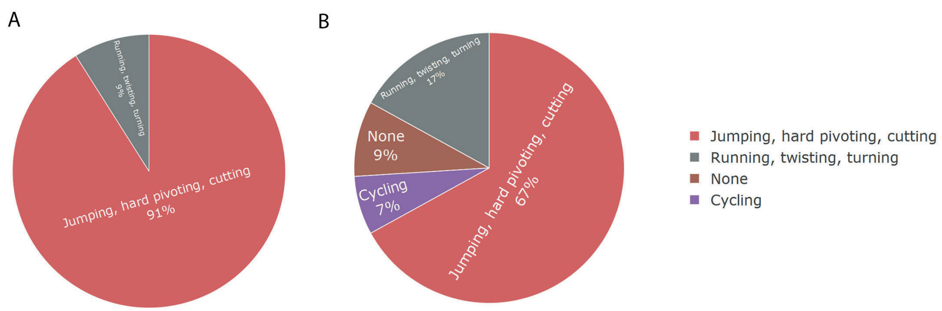


Figure 2.

Types of participated sports at pre-injury and 5-year follow-up based on level of physical demand(A) Type of sports that participants playing at pre-injury; (B) Types of sports that participants playing at 5-year follow-up.

A total of 41 participants (71%) returned to their desired sport. Of these, 28 (68%) performed at or above their pre-injury level, while 13 (32%) performed below their pre-injury level. A total of 12 participants (21%) returned to participation by modifying their desired sports into other sports, whereas 5 (9%) completely ceased sports participation. The reasons for modifying or discontinuing sports participation were both PT-related and non-PT-related (TABLE 5).

The frequency of sports activity per week at the 5-year follow-up (median [IQR] 3 [2–3]) was significantly lower than the pre-injury level (3 [3–3]; $p < 0.001$). Specifically, 58% of participants reported a decrease in sports frequency, 6% reported an increase, and 36% maintained the same frequency.

Table 5. Sports Participation at 5-year Follow-up

Sports Participation	Values
How many times per week do you currently participate in sport?	3 [2-3]
How many hours do you practice your sport on average per week?	2 [2-4]
Category of sports participation	
Return to performance	28 (48)
Return to sport	13 (22)
Return to participation	12 (21)
Because of PT symptoms	3 (25)
Because of fear or re-injury of patellar tendon	2 (17)
Because of non-PT related reasons	7 (58)
Completely stopped sports participation	5 (9)
Because of PT symptoms	3 (60)
Because of non-PT related reasons	2 (40)

Values are median [interquartile range], n (%).

Abbreviation: PT: patellar tendinopathy

Subgroup Analysis of Non-Recovered Participants

In the subgroup analysis (TABLE 6), participants who felt unrecovered had greater impairments in pain levels, disability, quality of life, and sports participation at 5 years compared to those who felt recovered (all $P < .05$).

Table 6. Subgroup Analysis of Participants with Self-Perceived Recovery and Non-Recovery for the Patient-Reported Outcomes at 5 Years[a]

Outcomes at 5 years	Recovery (n=44)	Non-recovery (n=14)	P value
Pain levels during ADL	1 [0-2]	4 [2-5]	<0.001
Pain levels during sports[b]	1 [0-3]	6 [5-7]	<0.001
VISA-P score	90 [75-87]	73 [63-78]	<0.001
EQ-5D-3L index	0.987 [0.984-0.987]	0.898 [0.846-0.984]	0.002
Sports participation, n(%)			0.027
Return to performance	25 (57)	3 (21)	
Return to sport	6 (14)	7 (50)	
Return to participation	9 (20)	3 (21)	
Quit sports	4 (9)	1 (7)	
Frequency of sports activity	3 [3-4]	3 [2-3]	<0.001
Duration of sports activity	2 [2-4]	3 [2-3]	0.926

[a]The values are denoted as the median [interquartile range, IQR] except sports participation.

[b]A total of 53 participants rated their pain levels due to 5 having quit sports at 5 years.

Abbreviations: VISA-P: the Victorian Institute of Sport Assessment-Patellar; VAS: Visual Analogue Scale; EQ-5D-3L: European Quality of Life-3 Dimensions.

Prognostic Factors

In the univariable analyses, higher BMI and VISA-P at baseline were sufficiently associated with a higher probability of self-perceived recovery to be entered into the multivariable model. However, none were statistically significant in the multivariable analysis (TABLE 7).

Table 7. Univariable and Multivariable Models Evaluating the Association Between Baseline Factors and Self-Perceived Recovery Status[a]

	Univariatea	P value	Multivariatea	P value
Sex (female)	1.24 (0.40-7.61)	0.556		
Age, y	1.09 (0.93-1.30)	0.308		
BMI, kg/m^2	1.26 (0.97-1.74)	0.108[b]	1.27 (0.97-1.77)	0.112
Symptom Duration, wk	1.00 (0.99-1.00)	0.20		
VISA-P score, 0-100 point	1.03 (0.99-1.09)	0.148[b]	1.03 (0.99-1.09)	0.158
Quadriceps muscle strength, N/kg	0.95 (0.53-1.72)	0.862		

[a]Coefficients represent odds ratios (ORs).

[b]Factors with P-values were <.15 in the univariable model and subsequently entered into the multivariable analyses adjusted with treatment allocation (in our case, 2 variables were selected for inclusion in the multivariable analysis). Abbreviations: y: years; BMI: body mass index; wk: weeks; VISA-P: the Victorian Institute of Sport Assessment-Patellar; N/kg: Newton/kilograms.

Discussion

Our study provides long-term data for athletes with PT following current quality of care using comprehensive patient-reported outcomes. We observed self-perceived recovery rate of 76% and return to desired sport rate of 71%, along with sustained improvements in pain levels, disability, and quality of life. However, only 25% of participants reported complete relief from pain during sports at the 5-year follow-up, and mild yet persistent symptoms were observed in the majority of cases. Additionally, over half of the athletes reported reduced frequency in sports participation. Participants who felt unrecovered showed higher levels of pain and disability, lower quality of life and less likely to return to performance compared to those who experienced recovery. No prognostic factors for self-perceived recovery were identified.

Comparison of Current Findings with the Literature

Prognosis based on sports participation

To our knowledge, the long-term prognosis based on sports participation of PT has only been assessed in a limited number of small-scale studies. One earlier cohort study reported generally poor outcomes, with a sports cessation rate of 52% (9 of 17 athletes)[16] after 15 years. In contrast, our study found a substantially lower sports cessation rate of 9% (of these, 60% was because of PT symptoms), suggesting an improved prognosis based on perceived functional status. This overall improvement could be explained by advances in the management of PT over time[23]. Compared to this previous cohort, most of the athletes in our study followed a 24-week evidence-based[21] exercise regimen, such as progressive tendon-loading exercises or eccentric exercise therapy, and often supplemented with active or passive therapeutic modalities. These approaches likely contributed to a higher return to sport rate. This speculation was also supported by a recent cohort study showing that 57% (16 of 28 male athletes) returned to their preinjury sport at 3-4 years following a structured 12-week exercise treatment[1]. However, caution is warranted when interpreting these differences, given the variations in outcome measures and study populations. Previous studies often employed dichotomous outcomes (e.g., quit or not), whereas our study adhered to the current consensus guidelines[2] by categorizing sports participation into four distinct levels (returning to performance, sport, participation and quitting). This approach allowed us to provide a more nuanced and comprehensive assessment of sports participation, reducing the potential overestimation associated with dichotomous measures, particularly in studies with small sample sizes. Furthermore, differences in cohort characteristics may account for some observed discrepancies. For example, 2 prior studies focused exclusively on male athletes[1,16], and another study limited to volleyball players[36]. In contrast, our study included female participants and a boarder range of sports. Additionally, the longer follow-up in previous studies may have contributed to the higher sports cessation rate.

Prognosis based on global rating of change (GROC)

To our knowledge, no study in PT has reported a long-term prognosis based on the outcome measured by GROC. However, a favorable prognosis—79% of patients felt recovered based on a dichotomized GROC—has been reported after 1 year of education and exercise therapy in patients with gluteal tendinopathy[22]. This instrument has been recognized as one of the core outcome measures in tendinopathy research[11,35]. Since prognosis is a major concern for patients[13], integrating patient-centered outcome measures such as a GROC is highly relevant. The GROC allows individuals to focus on the concerns most relevant to them[15], providing a holistic view when patients rate their course of conditions. Consistent with findings in other musculoskeletal disease studies[15,34], we found participants who felt recovered had larger improvements in pain levels, disability and quality of life (Appendix B, Table B2), indicating the GROC can reflect a multifaceted perspective on perceived change. Moreover, we observed a higher percentage of participants felt recovered (76%) compared to those achieving the MCID for domain-specific measures (e.g., 61-67% for pain levels during ADL and sports, and 71% for VISA-P score). This finding is not surprising as GROC has been reported to include additional constructs[15], including aspects like change in emotional well-being[17], that are not fully reflected in these specific clinical instruments. One notable finding from our study is that the majority of participants felt recovered despite the persistence of pain (25% reported complete relief from pain). It seems plausible that PT athletes may care more about overall improvement during recovery than about complete absence of pain.

Clinical Implications

This observation has important clinical implications. Specifically, healthcare providers could use this data to facilitate patient education and decision-making while administering treatment.

Healthcare professionals should communicate this generally positive prognosis to patients, underscoring that most athletes with PT experience recovery and return to their pre-injury sport under current standard treatment approaches. It is important to balance this optimism with realistic expectations about the possibility of mild but long-standing symptoms. By addressing this, patients may remain motivated and engaged in their rehabilitation programs rather than being frustrated if full symptom resolution is not achieved.

It is also crucial to recognize that approximately one-quarter of athletes in our study did not feel recovered, highlighting the need for a more comprehensive rehabilitation approach. Beyond current therapies, these patients may benefit from an extensive intervention that integrates effective pain management[23], functional maintenance, and providing psychological or social support[32]. Such approaches may also help athletes overcome the fear of re-injury, facilitating a fully return to sport. More invasive treatments, like surgery[3], could be considered for refractory conditions.

Strengths and Limitations

We are the first and largest to report 5-year follow-up data in athletes with PT using a broad spectrum of outcome measures aligned with the core domains established for tendinopathy research[11,35]. Additionally, we were able to contact a high number of participants during follow-up and we did not identify relevant differences between responders and non-responders, which increases the likelihood that responders were representative of the included group at baseline. The relationship between GROC and outcome changes assessed by serial measure instruments was examined. Thus, the potential recall bias of using GROC was reduced as much as possible.

However, several limitations of this study should also be acknowledged. The generalizability of our findings may be limited, as participants were drawn from a randomized trial with strict inclusion criteria, potentially excluding individuals relevant to the broader prognosis[26,29]. In addition, our findings could only reflect the average prognosis of people managed with the aforementioned treatment modalities. These results may not necessarily apply to settings with other quality of care. For example, the exercise regimens used in the original trial may differ from those typically applied in general clinical practice. Another limitation of this study is the difference in pain measurement tools used at baseline (Visual Analogue Scale) and during follow-up (Numerical Rating Scale via digital questionnaire). However, both utilized a 0–10 scale and are generally considered interchangeable in tendinopathy practice[8]. Furthermore, we were unable to further explore the variability of the prognosis based on different treatment strategies, such as different exercise programs (PTLE versus EET), conservative versus surgical (only one participant received surgery), and passive treatments versus injections, due to the limited sample size and the application of co-interventions in most participants. We did not ask participants the time frames when they returned to sport. Considering the longstanding nature of PT, only reporting this outcome as a binary aspect does not fully cover the course of this outcome. Finally, we did not identify any relevant prognostic factors due to the limited statistical power, which may hamper us in identifying potential targets to improve prognosis or treatment effectiveness. Larger cohort studies with extended follow-up periods are warranted. We used the categorization of sports participation to present prognosis in sports activity. Future research could address this using real-time GPS-based data to better capture the changes in activity level and accurately observe the influence of patellar tendinopathy.

Conclusion

Athletes with PT have a generally acceptable long-term prognosis following physical therapy, with the majority receiving multiple interventions. Approximately three-quarters of the patients feel recovered and returned to their desired sports at 5 years follow-up, with significant improvement in specific outcome domains. However, it should be noted that only 25% of the athletes who maintain sports participation are completely pain free, and 58% of the cases reduced frequency in sports activity. Nearly a quarter of participants did not feel recovered and perceived higher pain levels and disability, lower quality of life and less likely to return to performance. No prognostic factors for recovery could be identified. Clinicians may use these findings to estimate the average prognosis when educating athletes with PT.

References

1. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, Magnusson SP. Clinical Outcomes and Tendon Structure at 3- to 4-Year Follow-up After Exercise-Based Treatment of Patellar Tendinopathy: A Prospective Study. *Orthopaedic Journal of Sports Medicine*. 2024;12(10):23259671241280192; PMID:39430115
2. Ardern CL, Glasgow P, Schneiders A, et al. 2016 Consensus statement on return to sport from the First World Congress in Sports Physical Therapy, Bern. *Br J Sports Med*. 2016;50(14):853-864; PMID:27226389
3. Arner JW, Kaeding CC, Bradley JP. Management of Patellar Tendinopathy. *Arthroscopy*. 2024;40(1):13-15; PMID:38123261
4. Barber-Westin S, Noyes FR. Assessment of Sports Participation Levels Following Knee Injuries. *Sports Medicine*. 1999/07/01 1999;28(1):1-10;
5. Breda SJ, Oei EHG, Zwerver J, et al. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. *Br J Sports Med*. 2021;55(9):501-509;
6. Challoumas D, Pedret C, Biddle M, et al. Management of patellar tendinopathy: a systematic review and network meta-analysis of randomised studies. *BMJ Open Sport Exerc Med*. 2021;7(4):e001110;
7. Challoumas D, Zouvani A, Creavin K, et al. Determining minimal important differences for patient-reported outcome measures in shoulder, lateral elbow, patellar and Achilles tendinopathies using distribution-based methods. *BMC Musculoskelet Disord*. 2023;24(1):158; PMID:36864412
8. de Vos R-J, Gravare Silbernagel K, Malliaras P, et al. ICON 2023: International Scientific Tendinopathy Symposium Consensus - the core outcome set for Achilles tendinopathy (COS-AT) using a systematic review and a Delphi study of professional participants and patients. *Br J Sports Med*. 2024; PMID:39271248
9. De Vries AJ, Koolhaas W, Zwerver J, et al. The impact of patellar tendinopathy on sports and work performance in active athletes. *Res Sports Med*. 2017;25(3):253-265; PMID:28391723
10. Deng J, Breda SJ, Eygendaal D, Oei EH, de Vos R-J. Association Between Physical Tests and Patients-Reported Outcomes in Athletes Performing Exercise Therapy for Patellar Tendinopathy: A Secondary Analysis of the JUMPER Study. *Am J Sports Med*. 2023;51(13):3523-3532; PMID:37815096
11. Fearon AM, Grimaldi A, Mellor R, et al. ICON 2020-International Scientific Tendinopathy Symposium Consensus: the development of a core outcome set for gluteal tendinopathy. *Br J Sports Med*. 2024;58(5):245-254; PMID:38216320
12. Fernandez-Jaén T, Sanz-Zapata F, Cortés JM, et al. Proposal for a Clinical Analysis of Patellar Tendon Pathologies: In Search of Efficient Therapeutic Indications. *Orthopaedic Journal of Sports Medicine*. 2020;8(8):2325967120946312; PMID:32944587
13. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ*. 2013;346:e5595; PMID:23386360
14. Hernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. Mar 2014;48(6):453-7; PMID:23012320
15. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *The Journal of manual & manipulative therapy*. 2009;17(3):163-170; PMID:20046623
16. Kettunen JA, Kvist M, Alanen E, Kujala UM. Long-term prognosis for jumper's knee in male athletes. A prospective follow-up study. *Am J Sports Med*. 2002;30(5):689-692; PMID:12239003

17. Kounali D, Button KS, Lewis G, et al. How much change is enough? Evidence from a longitudinal study on depression in UK primary care. *Psychol Med*. 2022;52(10):1875-1882; PMID:33138872
18. Lankhorst NE, van Middelkoop M, Crossley KM, et al. Factors that predict a poor outcome 5-8 years after the diagnosis of patellofemoral pain: a multicentre observational analysis. *Br J Sports Med*. 2016;50(14):881-886; PMID:26463119
19. Lian OB, Engebretsen L, Bahr R. Prevalence of jumper's knee among elite athletes from different sports: a cross-sectional study. *Am J Sports Med*. 2005;33(4):561-567; PMID:15722279
20. M Versteegh M, M Vermeulen K, M A A Evers S, de Wit GA, Prenger R, A Stolk E. Dutch Tariff for the Five-Level Version of EQ-5D. *Value Health*. 2016;19(4):343-352; PMID:27325326
21. Marigi EM, Buckley P, Razi F, et al. Patellar Tendinopathy: Critical Analysis Review of Current Nonoperative Treatments. *JBJS Rev*. 2022;10(3) PMID:35358114
22. Mellor R, Bennell K, Grimaldi A, et al. Education plus exercise versus corticosteroid injection use versus a wait and see approach on global outcome and pain from gluteal tendinopathy: prospective, single blinded, randomised clinical trial. *BMJ*. 2018;361:k1662;
23. Millar NL, Silbernagel KG, Thorborg K, et al. Tendinopathy. *Nat Rev Dis Primers*. 2021/01/07 2021;7(1):1;
24. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009;338:b375; PMID:19237405
25. Nutarelli S, da Lodi CMT, Cook JL, Deabate L, Filardo G. Epidemiology of Patellar Tendinopathy in Athletes and the General Population: A Systematic Review and Meta-analysis. *Orthopaedic Journal of Sports Medicine*. 2023;11(6):23259671231173659; PMID:37347023
26. Pajouheshnia R, Groenwold RHH, Peelen LM, Reitsma JB, Moons KGM. When and how to use data from randomised trials to develop or validate prognostic models. *BMJ*. 2019;365:l2154; PMID:31142454
27. Rathleff MS, Winiarski L, Krommes K, et al. Pain, Sports Participation, and Physical Function in Adolescents With Patellofemoral Pain and Osgood-Schlatter Disease: A Matched Cross-sectional Study. *J Orthop Sports Phys Ther*. 2020;50(3):149-157; PMID:31905093
28. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med*. 2013;10(2):e1001380; PMID:23393429
29. Riley RD, Moons KGM, Snell KIE, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ*. 2019;364:k4597; PMID:30700442
30. Rio EK, Mc Auliffe S, Kuipers I, et al. ICON PART-T 2019-International Scientific Tendinopathy Symposium Consensus: recommended standards for reporting participant characteristics in tendinopathy research (PART-T). *Br J Sports Med*. 2020;54(11):627-630; PMID:31519545
31. Rudavsky A, Cook J. Physiotherapy management of patellar tendinopathy (jumper's knee). *J Physiother*. 2014;60(3):122-129; PMID:25092419
32. Stubbs C, McAuliffe S, Chimenti RL, et al. Which Psychological and Psychosocial Constructs Are Important to Measure in Future Tendinopathy Clinical Trials? A Modified International Delphi Study With Expert Clinician/Researchers and People With Tendinopathy. *J Orthop Sports Phys Ther*. 2024;54(1):14-25; PMID:37729020
33. van Linschoten R, van Middelkoop M, Berger MY, et al. Supervised exercise therapy versus usual care for patellofemoral pain syndrome: an open label randomised controlled trial. *BMJ*. 2009;339:b4074; PMID:19843565
34. Vasconcelos GS, Andreucci A, Vicenzino B, et al. Global Rating of Change for Better or Worse – what does it mean when patients who are treated for patellofemoral pain rate their change? *Journal of Orthopaedic & Sports Physical Therapy*. 2024;0(ja):1-22;

35. Vicenzino B, de Vos R-J, Alfredson H, et al. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. *Br J Sports Med*. 2020;54(8):444-451;
36. Visnes H, Bache-Mathiesen LK, Yamaguchi T, et al. Long-term Prognosis of Patellar Tendinopathy (Jumper's Knee) in Young, Elite Volleyball Players: Tendon Changes 11 Years After Baseline. *Am J Sports Med*. 2024:3635465241284648; PMID:39439271
37. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*. 2008;61(4):344-349; PMID:18313558
38. Zwerver J, Kramer T, van den Akker-Scheek I. Validity and reliability of the Dutch translation of the VISA-P questionnaire for patellar tendinopathy. *BMC Musculoskelet Disord*. 2009;10:102; PMID:19671174

Supplementary material

Appendix A online questionnaire

Appendix B:

Table B1 Comparison in baseline variables between responders and non-responders

Table B2 Comparison of Improvement in Pain Level, Disability and Quality of Life in Participants with Self-Perceived Recovery and Non-Recovery

Self-rating recovery

Question: Compared to your complaints at the start of your participation in the JUMPER study, how much has your condition recovered?

1. Completely recovered
2. Significantly improved
3. Slightly improved
4. Remained the same
5. Slightly worsened
6. Significantly worsened
7. Worse than ever

Sports participation

Question: Are you currently practicing your desired sport that you performed before suffering your patellar tendon injury?

1. Yes
2. No

Question: are you practicing your desired sport from before your patellar tendon injury at your former level?

1. Yes
2. No, I am performing lower than my preinjury level
3. No, I am performing better than my preinjury level

Question: If you did not return to your desired sports, did you modify your activities or make a transition to playing other sports?

1. Yes
2. No

Question: If you modified or changed your desired sports, what's your current sport?

Question: Are you currently not participating in any sports?

1. Yes
2. No

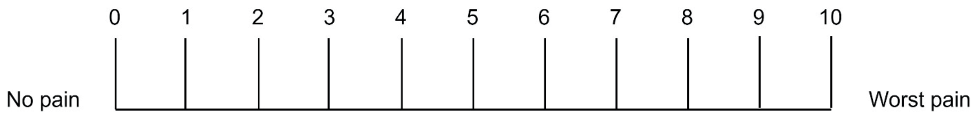
Question: Did you modify your current sports activities or stop sports participation because of the following reasons? (please rank)

1. Patellar tendon injury
2. Fear of re-injury of my patellar tendon
3. Change of interest
4. Lack of time
5. COVID-19 restrictions
6. Other injuries
7. Others

Question: How often do you practice your sport per week?

Question: How many hours do you practice your sport on average per week?

Question: Indicate the extent of pain from your patellar tendon injury during activities of daily living (ADL) and your recent sport activity respectively in the past week. Choose the applicable answer.



Health care consumption

Question: This question is about the treatments you (possibly) applied for your patellar tendon injury between 6 months after the start of your participation in the JUMPER study and now (total of 54 months). Which treatments were applied to promote the recovery of the patellar tendon injury? (multiple answers possible)

1. None
2. Rest
3. Adjusting sports activities
4. Strengthening the muscles of the lower limbs
5. Stretching the muscles of the lower limbs
6. Foot orthoses (orthotic insoles, adjusted shoes)
7. Knee brace
8. Patellar strap

9. Medical taping
10. Massage therapy
11. Manual therapy
12. Therapeutic ultrasound
13. Shock-wave therapy
14. Platelet-rich Plasma (PRP)
15. Corticosteroid injection
16. Prolotherapy
17. Dry needling
18. Percutaneous needle electrolysis
19. Medication (use of paracetamol, anti-inflammatory agents)
20. Surgery
21. Others, namely:

Validated questionnaires in the Dutch language:

VISA-P score

EQ-5D-3L

Table B1: The comparison between responders and non-responders in baseline characteristics[a]

Characteristics	All (n=76)	Responders (n=58)	Non-responders (n=18)	P value[b]
Demographic				
Age, y	25 (4)	25 (4)	24 (4)	0.709
Gender, male, n (%)	58 (76)	42 (72)	16 (89)	0.211
Height, cm	185 (9)	185 (10)	185 (8)	0.937
BMI, Kg/m^2	23.7 [22.0-25.0]	23.5 [22.0-25.1]	23.8 [22.4-24.8]	0.625
Tendinopathy descriptors				
Symptom duration, wk	104 [49-208]	106 [52-247]	75 [37-104]	0.081
Bilateral tendon involvement, n(%)	32 (42)	27 (47)	5 (28)	0.183
VISA-P score, (0-100)	55 (13)	56 (13)	54 (13)	0.659
VAS during a single-leg decline squat, (0-10)	5 [3-7]	5 [3-7]	6 [3-6]	0.763
Sports characteristics before injury, n(%)				
Sports activities by CSAS, n(%)				0.747
Level 1: 4-7 d/wk	17 (22)	14 (24)	3 (17)	
Level 2: 1-3 d/wk	59 (78)	44 (76)	15 (83)	
Clinical assessment				
Quadriceps muscle strength, N/kg	4.7 (1.0)	4.7 (1.0)	4.7 (1.0)	0.916
EQ-5D-3L index (0-1)	0.84 [0.81-0.84]	0.84 [0.81-0.84]	0.84 [0.81-0.90]	0.486

[a]Data are reported mean (SD) and median [IQR]. [b]For continuous data, a t-test was used to compare the two groups. Otherwise, a Mann-Whitney U was performed. For categorical data, we used Fisher's exact test. Abbreviations: y: years; BMI: body mass index; wk: weeks; N/kg: Newton/kilogram; CSAS: Cincinnati Sports Activity scale; VAS: visual analogue scale; VISA-P: the Victorian Institute of Sport Assessment-Patellar; EQ-5D-3L: European Quality of Life-3 Dimensions.

Table B2: Comparison of improvement in pain level, disability and quality of life in participants with self-perceived recovery and non-recovery[a]

	Recovery (n=44)	Non-recovery (n=14)	P value
Improvement in VAS during ADL	3 [2-5]	1 [-2-2]	.002
Improvement in VAS during sports	5 [4-7]	2 [-1-3]	<.001
Improvement in VISA-P score	29 [20-40]	13 [2-23]	.017
Improvement in EQ-5D-3L index	0.144 [0.057-0.176]	0.012 [-0.016-0.093]	.003

[a]The values are denoted as the absolute median [interquartile range, IQR] by subtracting the value from 5 years to baseline. Abbreviations: VISA-P: the Victorian Institute of Sport Assessment-Patellar; VAS: visual analogue scale; EQ-5D-3L: European Quality of Life-3 Dimensions.

4

Chapter 4

Do Physical or Imaging Changes Explain The Effectiveness Of Progressive Tendon Loading Exercises? A causal mediation analysis of athletes with patellar tendinopathy

Jie Deng, Jos Runhaar, Stephan J Breda, Edwin HG Oei, Denise Eygendaal, Robert-Jan de Vos

J Sci Med Sport. 2025 Jun;28(6):458-464.

Abstract

Objectives: To investigate whether the effectiveness of progressive tendon loading exercises (PTLE) on patellar tendinopathy is mediated through changes in physical or imaging properties.

Design: Mediation analyses based on a randomized clinical trial (n=76) in patellar tendinopathy comparing PTLE with eccentric exercise therapy (EET).

Methods: Pain-related disability on Victorian Institute of Sports Assessment–Patella (VISA-P, 0-100) and pain after single-leg decline squat (VAS-SLDS, 0-10) at 24 weeks were outcome measures. Selected mediators, including the physical (quadriceps muscle strength, ankle dorsiflexion range, jumping performance) and imaging domains (ultrasonographic tendon thickness and degree of neovascularization), were measured at 12 weeks. Directed acyclic graphs were performed to identify critical confounders. Causal mediation analysis was used to estimate natural indirect, natural direct and total effects by a simulation approach under the counterfactual framework.

Results: Complete data from 61 of 76 participants were included. There was no evidence showing that the beneficial effect of PTLE on VISA-P or VAS-SLDS outcomes was mediated by changes in any of the selected physical or imaging variables. The indirect effects for all mediators were unsubstantial (estimates ranging from -1.63 to 1.53 on VISA-P and -0.20 to 0.19 on VAS-SLDS), with all 95% confidence intervals containing zero.

Conclusions: The beneficial effect of PTLE on patellar tendinopathy was not mediated by changes in physical properties, tendon thickness or degree of neovascularization. Healthcare professionals may consider exploring other potential factors when managing patients with patellar tendinopathy, but further large-scale research is needed to confirm these results and to identify alternative treatment targets.

Introduction

Patellar tendinopathy (PT) is a highly prevalent injury in jumping athletes[1], leading to longstanding pain, decreased work productivity, and impaired sports performance[2]. Therapeutic exercises remain the cornerstone treatment for PT[3, 4]. We recently found that a 4-stage progressive tendon loading exercise therapy (PTLE) over 24 weeks leads to more improved symptom severity when compared to eccentric exercise therapy (EET)[5].

Causal mediation analyses can offer a robust methodological framework to elucidate how an intervention works by identifying potential targets or mediators for an intervention that could affect the outcome[6, 7]. Applying this approach for our trial can help advance theoretical understanding and optimize PTLE to improve effectiveness. PTLE, which incorporates isometric, isotonic, energy storage and sport-specific exercises, was designed to gradually load the patellar tendon, theoretically allowing tendons to increase tolerance to load[8]. However, it still needs to be determined whether the effects of this treatment are mediated by the improvement in physical properties (e.g., muscle strength). Additionally, from a clinical perspective, structural normalization, such as reduction in tendon thickening[9] and intratendinous neovascularization[10], is thought to be one of the main reasons for the effectiveness of exercise therapy. Nevertheless, evidence on this underlying mechanism by which the PTLE may exert its effects also remains scarce.

This study aimed to determine whether the effects of PTLE compared to EET on disability and pain at 24 weeks were mediated by changes in physical or ultrasound-based imaging properties at 12 weeks. We hypothesized that PTLE effects may be mediated through changes in quadriceps muscle strength, ankle dorsiflexion range, jumping performance, tendon thickness, and degree of neovascularization.

Methods

Design and Participants

This mediation analysis is reported following the AGReMA Statement[7], and is a secondary analysis from a previous published randomized control trial (JUMPER study) comparing the effectiveness between PTLE and EET in athletes with PT[5]. This RCT showed that PTLE resulted in a significantly larger improvement in disability over 24 weeks when compared to EET[5]. From January 2017 to July 2019, participants were eligible if they were aged between 18 and 35 years, participated in sports at least three times a week, and reported less than 80 of 100 points of the Victorian Institute of Sports Assessment–Patella (VISA-P) questionnaire. The diagnosis of PT was established based on clinical examination by a sports physician, and was confirmed on ultrasound performed by a radiologist. Further details on diagnostic criteria, inclusion and exclusion criteria have been reported elsewhere[5]. This trial was designed (ClinicalTrials.gov; NCT02938143) and conducted at Erasmus MC University Medical Center. The institutional review board approved the study (MEC-2016-500), and all participants provided written informed consent.

Interventions

PTLE - The program of PTLE incorporated isometric, isotonic, energy storage, and sport-specific exercises, monitored by regular assessment of pain response (Visual Analogue Scale, VAS) with a single-leg declined squat test (VAS-SLDS ≤3 points on a scale of 0–10). Participants initially performed daily single-leg isometric exercises in 60° knee flexion. A 70% maximal voluntary contraction load (5 repetitions of a 45-second hold) was recommended. As pain permitted, single-leg isotonic loading was performed every second day, initially within 10-60° of knee flexion and progressed to full knee extension, resistance corresponding from 15 repetitions maximum (RM) to 6RM. Energy-storage loading and sport-specific exercises were progressed and customized according to the individual's sport among our study population, mainly focusing on jumping and landing training.

EET - The control group comprised patients undergoing the EET program, and this program consisted of 2 stages: a pain-provoking single-leg squat on a decline board with a 25° slope (VAS ≥ 5/10) and sport-specific exercises.

These two unsupervised exercise therapies were delivered over 24 weeks. Details about the exercise protocols, education, and load management advice can be found in the previously published work[5].

Measurements

Outcome measures:

The outcomes of interest in this mediation analysis were pain-related disability assessed with the VISA-P questionnaire (0 to 100 points) and pain intensity after single-leg decline squat, rated on a VAS (VAS-SLDS) at 24 weeks. These outcome measures are in line with the core domains for tendinopathy[11].

Proposed mediators:

Potential mediators were collected at baseline and at 12 weeks after randomization. We pre-specified mediators using scientific evidence and clinical-perspective knowledge to avoid data dredging. The following mediators were selected: maximal isometric quadriceps muscle strength (N/kg); active ankle dorsiflexion range (degrees); jumping performance (cm); ultrasonographic tendon thickness (mm) and tendon neovascularization (low / high) using power Doppler. Details of measurement for each mediator can be found in Table 1 and Appendix A.

Confounders:

It is reasonable to assume that no confounders between exposure (intervention)-mediator and intervention-outcome effects were present due to the randomization process[6, 7]. For mediator-outcome effects, we composed directed acyclic graphs (DAGs) to identify potential confounders (Appendix A). This was modified based on previous literature[18] and peer feedback. Based on these DAGs, we utilized the online tool 'DAGitty' (https://dagitty.net/) to identify the minimal set of confounders that should be adjusted for the mediator-outcome relationship, which included age, sex, physical activity level, as well as respective baseline values for the mediators studied (Fig.1).

Statistical analysis:

We conducted causal mediation analyses using the simulation-based approach under the counterfactual framework[6]. All analyses were conducted using R studio version 2023.12.0 with 'mediation' package. All codes relevant to this study can be found in the Supplementary file. Complete case analyses (participants with complete data on outcomes, mediators and confounders) were conducted.

Table 1. The Description of the Rationale and Measurement for Mediators within Assumed Causal Relationships

Mediators	Rationale	Measurement[a]
Knee muscle strength	Individuals with patellar tendinopathy have significantly lower isometric and concentric knee extensor strength than those with asymptomatic controls[12]. PTLE was targeted to improve knee muscle strength by gradually increasing the load on the patellar tendon over time, allowing the tendon to adapt in response to the increasing demands. Together, this could improve the load tolerance assessed by VAS-SLDS. Increased knee function can also contribute to an improved VISA-P score.	Isometric quadriceps muscle strength by a handheld dynamometer.
Ankle dorsiflexion range	Lower ankle dorsiflexion during landing has been found to be associated with patellar tendinopathy[13]. The PTLE protocol was developed to augment the capacity for active ankle dorsiflexion range of motion during jump-landing maneuvers, such as drop landings, with the goal of enhancing athletic performance, which could improve load tolerance. Furthermore, the ability to optimize ankle movement could be important for lower limb control, which could positively influence perceived symptoms and function.	The ankle dorsiflexion range (or flexibility of the soleus muscle) was tested by using a weight-bearing dorsiflexion lunge test.
Jumping performance	PTLE, especially its energy-storage stage, aimed to increase jumping performance through gradual jumping training. This is likely to improve patients' performance. In addition, improved jumping ability is highly associated with increased load tolerance[8], which may also affect the overall knee symptoms and function.	Vertical jump height was measured in centimeters using a digital vertical jump meter.
Tendon Thickness	Increased tendon thickness or thickening is a morphological feature of patellar tendinopathy. This alteration may be predictive of symptom development[14] and could potentially be associated with less favorable outcomes by conservative treatment[15] in patellar tendinopathy. A reduction in thickness after a 12-week heavy slow resistance training has been observed[9]. This adaption seems to be correlated with positive treatment outcomes in patellar tendinopathy[9, 16]. We hypothesized that decreased thickness in response to PTLE may partially explain the improvement in clinical outcomes.	The patellar tendon anteroposterior (AP) thickness was measured in the transverse plane at the thickest point within 1 cm distal to the inferior pole of the patella, using a built-in software involving caliper.

Table 1. Continued

Mediators	Rationale	Measurement[a]
Degree of neovascularization	Increased Doppler flow (frequently referred to as 'neovascularization') is a common ultrasonographic finding in tendinopathy. The vascular theory suggested that increased blood flow within the tendon may be associated with the persistence of symptoms in tendinopathy[17]. Research has shown that, after a heavy slow resistance in patellar tendinopathy, there is an improved VISA-P score and pain, and a concomitant reduction in neovascularization[9]. Based on this, we speculated that PTLE would induce a reduction in neovascularization that could partially explain the improvement in clinical outcomes.	Power Doppler was assessed by the modified Ohberg score (MOS), ranging from 0 to 4+ (0= no vessels; 1+= one vessel posterior to the patella tendon; 2+-4+ = 1,2,3,4 or more vessels throughout the tendon). We dichotomized this data into "no to low Doppler flow" (from 0 to 2+) and "moderate to high Doppler flow (from 3+ to 4+)".

Abbreviations: PTLE, progressive tendon loading exercises; EET, eccentric exercise therapy; VISA-P, Victorian Institute of Sports Assessment–Patella (VISA-P).

[a]Other details in each measurement are reported in the Appendix A.

4

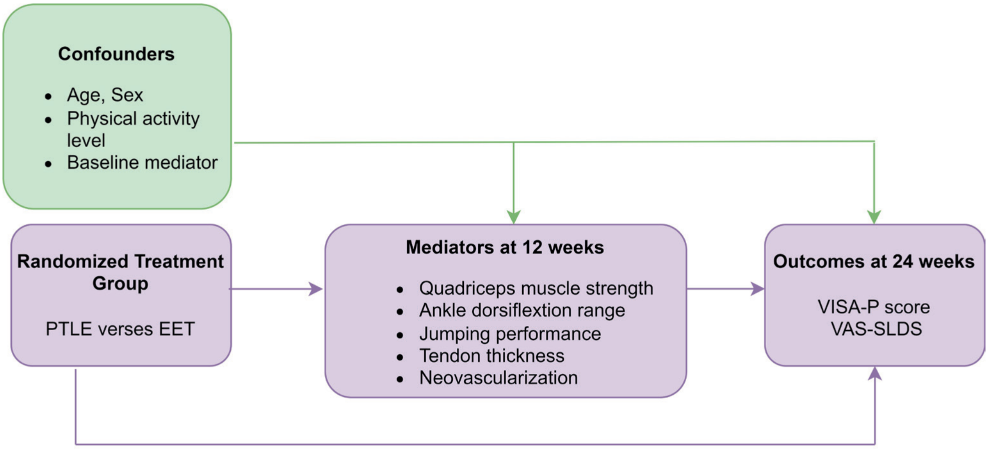


Figure 1. Single mediator model for VISA-P score and VAS-SLDS at 24 weeks.

Confounders (age, sex, physical activity level, mediator at baseline) which could distort the mediator-outcome relationship were adjusted.

In the primary assessment, we evaluated the strength of selected mediators for each outcome (VISA-P score and VAS-SLDS at 24 weeks), using single mediator models (Fig.1). For each model, we fitted two regression models: linear regression was used for the models on the outcomes, adjusted for treatment group, mediator at 12 weeks, and set of confounders mentioned above. The mediator model for neovascularization was logistic, while the models for physical variables and tendon thickness were linear. All models were adjusted for the treatment group and baseline mediator values. Assumptions for linear regression models (homogeneity of variance, normality of residuals and linearity) were checked using graphic methods. Linearity was not checked in the logistic regression analysis because there were no quantitative predictors. We assumed no interaction term between the exposure and mediator.

For each model, we estimated total effect (TE), natural direct effect (NDE), natural indirect effect (NIE), and proportion of effect mediated. The TE is the average effect of treatment on the outcome; the NIE is the average effect of treatment on the outcome through the selected mediating path. The NDE is the average effect of treatment on the outcome through other pathways. The proportion mediated is the ratio of NIE to TE. 95% Confidence intervals for these causal effects were generated by bootstrapped samples with 2000 simulations.

We conducted 2 sensitivity analyses:

1. Strong assumptions are required to obtain valid estimated for NIE and NDE. Specifically, there are no unmeasured confounders in exposure-mediator, exposure-outcome, and mediator-outcome relationships[6, 7] in mediation analyses. Thus, we employed the mediational E-value[19] to assess the robustness of these observed effects to potential unmeasured confounding. A large E-value suggests that only strong unmeasured confounding could nullify the observed association, while a small E-value indicates that even weak confounding could do so. Details in calculations can be found in Appendix A.
2. To assess the robustness of the influence of the missing data, we reanalyzed the data using 20 imputed datasets by multiple imputation given the missing mechanism is missing at random (MAR). For the imputation model, we included confounders described above as well as auxiliary variables (baseline BMI, quality of life [the EuroQol Group with three-level index, EQ-5D-3L index], VAS-SLDS and VISA-P score at 12 weeks). We applied the simulation approach for each imputed dataset to compute NIE and NDE, and subsequently combined them to obtain pooled point estimates and confidence intervals (CIs).

We had planned to apply multiple mediator models, but due to the absence of significant mediators in the single mediator models, this approach was not performed.

Results

Of the enrolled 76 participants in the original RCT, 9 (12%) were lost to follow-up at 24 weeks; 1 in the PTLE group and 8 in the EET group. Most of the participants (n=61, 80%) were included in the mediation analysis, contributing complete data on the treatment, mediator, outcome and confounders. The baseline characteristics for included and original participants can be found in Table 2.
The percentages of missing values per variable during follow-up were reported (Appendix A).

Table 2. The Description of Baseline Characteristics for Included and Original Datasets[a]

Baseline Characteristics	Included dataset (n = 61)	Original dataset (n = 76)
Age, y	24 (4)	25 (4)
Male, n(%)	44 (72)	58 (76)
BMI, kg/m^2	24.1 (3.0)	23.9 (2.9)
Physical activity[b], n (%)		
Level 1: 4-7 d/wk	13 (21)	17 (22)
Level 2: 1-3 d/wk	48 (79)	59 (78)
Symptom duration, wk	104 [52-260]	104 [49-208]
VAS-SLDS (0-10 points)	5 (2)	5 (2)
VISA-P score (0-100 points)	55 (13)	55 (13)
Treatment group, n (%)		
PTLE	34 (56)	38 (50)
EET	27 (44)	38 (50)
EQ-5D-3L index[c]	0.84 [0.80-0.84]	0.84 [0.81-0.84]
Quadriceps muscle strength, N/kg	4.6 (1.0)	4.7 (1.0)
Active ankle dorsiflexion range, (°)	42 (6)	41 (7)
Jumping performance, cm	49 (10)	50 (10)
Tendon AP thickness, mm	8.2 (2.2)	8.4 (2.3)
Neovascularization[d], n (%)		
No to low Doppler flow	21 (34)	25 (33)
Moderate to high Doppler flow	40 (66)	51 (67)

Abbreviations: BMI, body mass index; VAS-SLDS, visual analog scale after single-leg decline squat test; VISA-P, Victorian Institute of Sports Assessment–Patella; PTLE, progressive tendon loading exercises; EET, eccentric exercise therapy; EQ-5D-3L, a standardized measure of health status developed by the EuroQol Group with three-level version; AP, anterior-posterior; wk: weeks

[a]Values are mean (SD) if continuous data with normal distribution; otherwise, for continuous data with non-normal distribution, median [interquartile range, IQR] is used.

[b]Physical activity is semi-quantified by Cincinnati Sports Activity Scale (CSAS)[20].

[c]EQ-5D-3L index is calculated using a set of weights on Dutch population[21]. Values are anchored at 1 (full health) and 0 (a state as bad as being dead).

[d]No to low Doppler flow: from score 0 to score 2+, using modified Ohberg score [MOS]). Moderate to high Doppler flow: from score 3+ to score 4+.

Mediation Analysis

Table 3 displays the indirect, direct, and total effects for selected mediators on VISA-P score and VAS-SLDS at 24 weeks. Compared with EET, PTLE showed a superior mean improvement in VISA-P and reduction in VAS-SLDS at 24 weeks, which was supported by statistically significant NDE and TE. However, the selected mediators did not account for these beneficial outcomes, as indicated by the negligible indirect effects.

Table 3. Estimates of total, natural direct and indirect effects of treatment on VISA-P score and VAS-SLDS at 24 weeks for 5 potential mediators at 12 weeks[a]

	TE		NDE		NIE		Proportion mediated (95% CI)
	Effect (95% CI)	*p value*	Effect (95% CI)	*p value*	Effect (95% CI)	*p value*	
VISA-P score (0-100 points)							
Quadriceps muscle strength	11.75 (3.54 to 20.47)	0.005	11.18 (0.97 to 22.22)	0.032	0.57 (-4.45 to 4.84)	0.796	0.05 (-0.38 to 0.76)
Ankle dorsiflexion angle	12.14 (4.12 to 21.38)	0.004	12.30 (4.23 to 21.75)	0.004	-0.16 (-2.29 to 1.14)	0.797	NA[b]
Jumping performance	13.31 (5.04 to 21.93)	0.002	11.78 (3.83 to 20.43)	0.008	1.53 (-0.23 to 4.39)	0.100	0.12 (-0.02 to 0.39)
Patellar tendon thickness	12.30 (3.77 to 21.45)	0.002	12.82 (3.83 to 21.89)	0.005	-0.52 (-2.98 to 1.23)	0.599	NA[b]
Degree of neovascularization[c]	11.96 (3.98 to 22.19)	0.004	13.59 (3.67 to 23.32)	0.007	-1.63 (-4.18 to 1.48)	0.455	NA[b]
VAS-SLDS (0-10 points)							
Quadriceps muscle strength	-1.28 (-2.39 to -0.27)	0.018	-1.27 (-2.53 to -0.08)	0.034	-0.01 (-0.59 to 0.56)	0.955	0.01 (-0.66 to 0.72)
Ankle dorsiflexion angle	-1.35 (-2.41 to -0.31)	0.013	-1.42 (-2.49 to -0.40)	0.070	0.07 (-0.16 to 0.44)	0.611	NA[b]
Jumping performance	-1.36 (-2.44 to -0.35)	0.010	-1.17 (-2.30 to -0.11)	0.033	-0.20 (-0.51 to 0.05)	0.108	0.14 (-0.05 to 0.68)
Patellar tendon thickness	-1.19 (-2.24 to -0.15)	0.024	-1.24 (-2.33 to -0.12)	0.027	0.05 (-0.15 to 0.30)	0.705	NA[b]
Degree of neovascularization[c]	-1.14 (-2.38 to -0.11)	0.031	-1.33 (-2.52 to -0.11)	0.031	0.19 (-0.28 to 0.49)	0.565	NA[b]

[a]The point estimates and uncertainty estimates were measured as mean differences and 95% confidence intervals (CIs) with bootstrap samples. We used seed (set.seed) before running the mediation analysis to make results reproducible during simulation approach.

[b]The proportion mediated is only calculated when the TE and NIE have the same sign.

[c]The effect is denoted as log odds using logistic regression.

4

Sensitivity Analysis

Estimates of NDE and TE exhibited moderate robustness to unmeasured confounding in the exposure-outcome association, with E-values ranging from 2.74 to 3.75 for all estimates. The estimates of NIE were found to be robust to unmeasured confounding, with non-null E-values ranging from 1.24 to 1.93. Details of interpretation for these E-values in Appendix A.

Figure 2 presents the sensitivity analysis comparing datasets with complete cases to dataset with imputation. The magnitude of NDE in the imputed dataset was slightly smaller than that in the complete-case dataset. However, the positive effects on VISA-P score and VAS-SLDS at 24 weeks were still not mediated through these selected mediators in the imputed dataset. This implied our findings performed robustness to missing data.

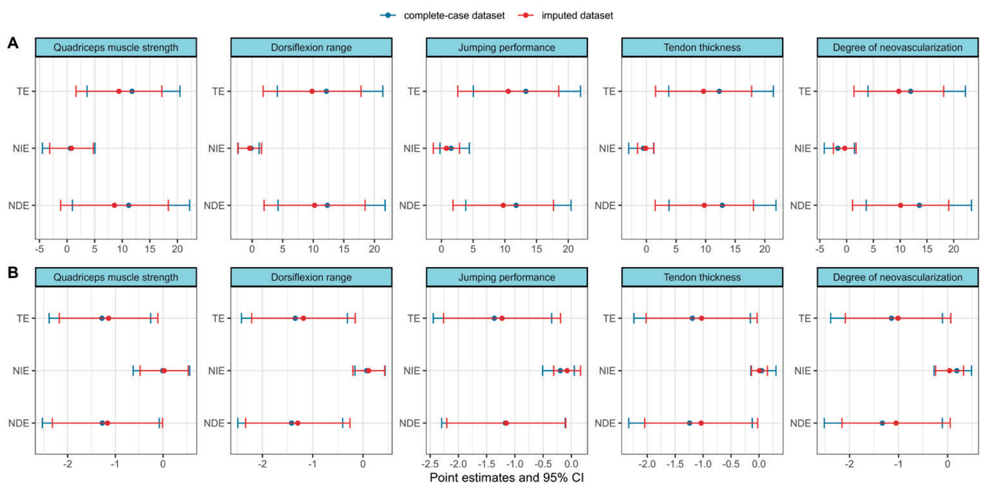


Figure. 2 Panel A and B show sensitivity analysis for causal effects on 24-week VISA-P, and VAS-SLDS respectively, comparing complete-case dataset (n=61) with imputed datasets (n=76). Dots represent point estimates. I bars indicates 95% confidence intervals by bootstrapped samples.

Discussion

This study is the first to use causal mediation analysis to investigate the potential underlying mechanisms behind the beneficial effect of PTLE in improving patellar tendinopathy. We aimed to examine whether the improvement in clinical outcomes at 24 weeks after PTLE compared with EET was through the change in physical and ultrasound-based imaging properties at 12 weeks. Our findings do not support our hypothesis that the improvement in disability and pain following PTLE is mediated through changes in quadriceps muscle strength, ankle dorsiflexion range, jumping performance, tendon thickness, or degree of neovascularization.

To date, PTLE shows promise as the initial conservative treatment for PT[5]. It is critical to understand how this intervention exerts its effect, which can advance and refine the theory to improve the effectiveness in clinical practice. Identifying mediators by conducting a mediation analysis from a trial is particularly relevant to answering this question[6].

Physical mediators

Physical properties remain pivotal in understanding the response to exercise treatment in musculoskeletal conditions[22, 23]. Previous mediation analyses performed in an exercise-based trial for gluteal tendinopathy[22], and patellofemoral pain[23]. Align with our findings, these authors also reported that improvements in pain and function, or the global rate of change were not explained by changes in muscle strength or joint movement. There is one study[24] suggesting that muscle strength acts as a weak mediating effect in knee osteoarthritis following exercise treatment, with around 2% of the effect mediated. Caution should be taken when comparing these results, considering that variables (intervention, mediator, and outcome) were diverse across these studies

One possibility for the absence of a mediating effect is that both exercise therapies were performed in a home-based and unsupervised way, which may have led to a less significant changes in these selected physical variables at 12 weeks between PTLE and EET (Appendix A). Another explanation could be that, compared to these easy-to-perform measurement, there are other physical variables that might be more sensitive and better at explaining changes in pain and function following PTLE. Several studies have reported altered muscle activation resulting from changes in the neuromuscular pathways in PT[25, 26]. A study with a small sample size indicated that immediate isometric contraction could reduce pain in patients with PT, parallel with a reduction in cortical inhibition[25]. These findings may provide new insight into neurophysiological mechanisms by which PTLE may work, as tolerable exercise-induced pain experience is the keystone for this specific loading program. This may be an area for future study.

Imaging mediators

Tendon thickening and neovascularization, as observed via ultrasound (US), are common features in PT[27]. These ultrasonographic findings substantiate the underlying pathological changes occurring in tendinopathy, such as increased water content, collagen disorganization as well as vascular and nerve ingrowth[28]. It is well established that healthy tendons respond favorably to mechanical loading or exercise, attributed to structural remodeling[28]. This adaptive capability underpins the hypothesis that the effectiveness of therapeutic exercises in PT may be achieved by targeting structural normalization, as observed in Achilles tendinopathy[29].

A systematic review by Drew et al[30] reported no association between improvement in pain and function and change in tendon thickness and neovascularization in tendinopathy. In contrast, another systematic review[10], specifically in PT and Achilles tendinopathy, implied a positive link. Nevertheless, it is crucial to emphasize that both studies only addressed associations rather than causal relationships due to the absence of mediation analysis in their methodologies. Our study addresses this gap in the field of PT.

We did not find any significant mediating effect in tendon thickness or the degree of neovascularization. The reason for this could be that the absolute between-group difference in these ultrasonographic findings are too subtle (e.g. less than 1 mm in tendon thickness) (Appendix A). A similar small effect on these conventional ultrasound findings were reported in patients with PT performing heavy slow resistance training (HSR) over EET[9]. These findings highlight the need for future research to explore other precise imaging variables that could accurately elucidate the mechanisms underlying the effectiveness of exercise. Advanced imaging techniques such as share-wave electrography (SWE)[31], ultrashort echo time (UTE) MRI[32] hold promise in this regard by offering more insights into the mechanical and material components. However, the practical application of these technologies in a clinical setting remains challenging due to their complexity and resource requirements.

Strengths and Limitations

This study was based on the largest RCT in the field of PT to date with longitudinal data, enabling a clear chronological order between exposure, mediator, and outcome, which is the premise for the direction of the causal relationship[7]. We further corroborated our findings through two sensitivity analyses by assessing the robustness of casual assumptions and the influence of missing data, which are crucial components in mediation analysis for validity. Measurement error was reduced as the same individual conducted all assessments, and the methods used for measurement demonstrated relatively high reliability.

One of the most important limitations of our study is that we did not collect psychological variables, such as pain catastrophizing, and pain self-efficacy. These factors could be potential mediators for PTLE over EET, as they appear to mediate the effectiveness of exercise therapy in other musculoskeletal conditions such as gluteal tendinopathy[22]. Since PTLE was designed to alleviate exercise-induced pain, investigating these pain-related psychological mechanisms would be plausible. Another limitation of this study is the small sample size (n=61), which may have limited power to detect stable indirect effects. This may also limit our ability to explore the interaction between a potential mediator and exposure or non-linear relationships in this analysis, given the speculation that the physical and structural response to exercise intervention could vary by patients. Future large-scale studies are needed to improve the stability of the mediation effects and to unravel these complex relationships.

Recommendation for future research

Future studies should consider investigating other potential mediators, such as neuromuscular adaptations or tendon stiffness, which may be potential targets for optimized interventions in patellar tendinopathy. Additionally, studies with sufficient sample size are essential to confirm our findings and to explore complex relationship between exercise effects and change in physical or imaging properties. Investigating long-term outcomes may also help to better understand the mechanisms of interventions.

Conclusion

The beneficial effect of progressive tendon loading exercises over eccentric exercise therapy in pain and pain-related disability for patellar tendinopathy is not mediated by changes in physical factors (quadriceps muscle strength, ankle dorsiflexion range, and jumping performance) or imaging factors (tendon thickness and degree of neovascularization). However, these findings should be confirmed in studies with larger sample size. We also emphasize the need to evaluate alternative mechanisms, as they may be important for clinicians to use as treatment targets.

References

1. Nutarelli S, da Lodi CMT, Cook JL, et al. Epidemiology of Patellar Tendinopathy in Athletes and the General Population: A Systematic Review and Meta-analysis. Orthop J Sports Med 2023;11(6):23259671231173659. http:// doi.org/10.1177/23259671231173659.
2. De Vries AJ, Koolhaas W, Zwerver J, et al. The impact of patellar tendinopathy on sports and work performance in active athletes. Research In Sports Medicine 2017;25(3):253-65. http:// doi.org/10.1080/15438627.2017.1314292.
3. Challoumas D, Crosbie G, O'Neill S, et al. Effectiveness of Exercise Treatments with or without Adjuncts for Common Lower Limb Tendinopathies: A Living Systematic Review and Network Meta-analysis. Sports Med Open 2023;9(1):71. http:// doi.org/10.1186/s40798-023-00616-1.
4. Millar NL, Silbernagel KG, Thorborg K, et al. Tendinopathy. Nat Rev Dis Primers 2021;7(1):1. http:// doi.org/10.1038/s41572-020-00234-1.
5. Breda SJ, Oei EHG, Zwerver J, et al. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. Br J Sports Med 2021;55(9):501-9. http:// doi.org/10.1136/bjsports-2020-103403.
6. Cashin AG, McAuley JH, VanderWeele TJ, Lee H. Understanding how health interventions or exposures produce their effects using mediation analysis. BMJ 2023;382:e071757. http:// doi.org/10.1136/bmj-2022-071757.
7. Lee H, Cashin AG, Lamb SE, et al. A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement. JAMA 2021;326(11):1045-56. http:// doi.org/10.1001/jama.2021.14075.
8. Malliaras P, Cook J, Purdam C, Rio E. Patellar Tendinopathy: Clinical Diagnosis, Load Management, and Advice for Challenging Case Presentations. J Orthop Sports Phys Ther 2015;45(11):887-98. http:// doi.org/10.2519/jospt.2015.5987.
9. Kongsgaard M, Kovanen V, Aagaard P, et al. Corticosteroid injections, eccentric decline squat training and heavy slow resistance training in patellar tendinopathy. Scand J Med Sci Sports 2009;19(6):790-802. http:// doi.org/10.1111/j.1600-0838.2009.00949.x.
10. Rabello LM, van den Akker-Scheek I, Brink MS, et al. Association Between Clinical and Imaging Outcomes After Therapeutic Loading Exercise in Patients Diagnosed With Achilles or Patellar Tendinopathy at Short- and Long-Term Follow-up: A Systematic Review. Clin J Sport Med 2020;30(4):390-403. http:// doi.org/10.1097/jsm.0000000000000624.
11. Vicenzino B, de Vos R-J, Alfredson H, et al. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. Br J Sports Med 2020;54(8):444-51. http:// doi.org/10.1136/bjsports-2019-100894.
12. Obst SJ, Peterson B, Heales LJ. Maximal Lower Limb Strength in Patellar Tendinopathy: A Systematic Review With Meta-Analysis. J Athl Train 2024;59(2):159-72. http:// doi.org/10.4085/1062-6050-0662.22.
13. Tayfur A, Haque A, Salles JI, et al. Are Landing Patterns in Jumping Athletes Associated with Patellar Tendinopathy? A Systematic Review with Evidence Gap Map and Meta-analysis. Sports Med 2022;52(1):123-37. http:// doi.org/10.1007/s40279-021-01550-6.
14. McAuliffe S, McCreesh K, Culloty F, et al. Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? A systematic review and meta-analysis. Br J Sports Med 2016;50(24):1516-23. http:// doi.org/10.1136/bjsports-2016-096288.

15. Golman M, Wright ML, Wong TT, et al. Rethinking Patellar Tendinopathy and Partial Patellar Tendon Tears: A Novel Classification System. Am J Sports Med 2020;48(2):359-69. http:// doi.org/10.1177/0363546519894333.
16. Beyer R, Kongsgaard M, Hougs Kjær B, et al. Heavy Slow Resistance Versus Eccentric Training as Treatment for Achilles Tendinopathy: A Randomized Controlled Trial. Am J Sports Med 2015;43(7):1704-11. http:// doi.org/10.1177/0363546515584760.
17. Pufe T, Petersen WJ, Mentlein R, Tillmann BN. The role of vasculature and angiogenesis for the pathogenesis of degenerative tendons disease. Scand J Med Sci Sports 2005;15(4):211-22.
18. van Rijn D, van den Akker-Scheek I, Steunebrink M, et al. Comparison of the Effect of 5 Different Treatment Options for Managing Patellar Tendinopathy: A Secondary Analysis. Clin J Sport Med 2019;29(3):181-7. http:// doi.org/10.1097/jsm.0000000000000520.
19. Smith LH, VanderWeele TJ. Mediational E-values: Approximate Sensitivity Analysis for Unmeasured Mediator–Outcome Confounding. Epidemiology 2019;30(6):835-7. http:// doi.org/10.1097/ede.0000000000001064.
20. Barber-Westin SD, Noyes FR. Assessment of sports participation levels following knee injuries. Sports Med 1999;28(1). http:// doi.org/10.2165/00007256-199928010-00001.
21. Janssen MF, Szende A, Cabases J, et al. Population norms for the EQ-5D-3L: a cross-country analysis of population surveys for 20 countries. Eur J Health Econ 2019;20(2):205-16. http:// doi.org/10.1007/s10198-018-0955-5.
22. Mellor R, Kasza J, Grimaldi A, et al. Mediators and Moderators of Education Plus Exercise on Perceived Improvement in Individuals With Gluteal Tendinopathy: An Exploratory Analysis of a 3-Arm Randomized Trial. J Orthop Sports Phys Ther 2022;52(12):826-36. http:// doi.org/10.2519/jospt.2022.11261.
23. Holden S, Matthews M, Rathleff MS, et al. How Do Hip Exercises Improve Pain in Individuals With Patellofemoral Pain? Secondary Mediation Analysis of Strength and Psychological Factors as Mechanisms. J Orthop Sports Phys Ther 2021;51(12):602-10. http:// doi.org/10.2519/jospt.2021.10674.
24. Runhaar J, Holden MA, Hattle M, et al. Mechanisms of action of therapeutic exercise for knee and hip OA remain a black box phenomenon: an individual patient data mediation study with the OA Trial Bank. RMD Open 2023;9(3). http:// doi.org/10.1136/rmdopen-2023-003220.
25. Rio E, Kidgell D, Purdam C, et al. Isometric exercise induces analgesia and reduces inhibition in patellar tendinopathy. Br J Sports Med 2015;49(19):1277-83. http:// doi.org/10.1136/bjsports-2014-094386.
26. Sprague AL, Couppé C, Pohlig RT, et al. Relationships between tendon structure and clinical impairments in patients with patellar tendinopathy. J Orthop Res 2022. http:// doi.org/10.1002/jor.25262.
27. Docking SI, Ooi CC, Connell D. Tendinopathy: Is Imaging Telling Us the Entire Story? J Orthop Sports Phys Ther 2015;45(11):842-52. http:// doi.org/10.2519/jospt.2015.5880.
28. Magnusson SP, Kjaer M. The impact of loading, unloading, ageing and injury on the human tendon. J Physiol 2019;597(5):1283-98. http:// doi.org/10.1113/jp275450.
29. de Jonge S, Tol JL, Weir A, et al. The Tendon Structure Returns to Asymptomatic Values in Nonoperatively Treated Achilles Tendinopathy but Is Not Associated With Symptoms: A Prospective Study. Am J Sports Med 2015;43(12):2950-8. http:// doi.org/10.1177/0363546515605077.
30. Drew BT, Smith TO, Littlewood C, Sturrock B. Do structural changes (eg, collagen/matrix) explain the response to therapeutic exercises in tendinopathy: a systematic review. Br J Sports Med 2014;48(12):966-72. http:// doi.org/10.1136/bjsports-2012-091285.

31. Breda SJ, van der Vlist A, de Vos R-J, et al. The association between patellar tendon stiffness measured with shear-wave elastography and patellar tendinopathy—a case-control study. Eur Radiol 2020;30(11):5942-51. http:// doi.org/10.1007/s00330-020-06952-0.
32. Breda SJ, de Vos R-J, Poot DHJ, et al. Association Between T Relaxation Times Derived From Ultrashort Echo Time MRI and Symptoms During Exercise Therapy for Patellar Tendinopathy: A Large Prospective Study. J Magn Reson Imaging 2021;54(5):1596-605. http:// doi.org/10.1002/jmri.27751.

Supplementary material

Supplementary material for this paper can be found online:

https://doi.org/10.1016/j.jsams.2024.12.006

5

Chapter 5

Results of Physical Tests Have No Prognostic Value for Patient-reported Outcomes in Athletes Performing Exercise Therapy for Patellar Tendinopathy: A secondary analysis of JUMPER study

Jie Deng, Stephan J Breda, Denise Eygendaal, Edwin HG Oei, Robert-Jan de Vos

Am J Sports Med. 2023 Nov;51(13):3523-3532.

Abstract

Background: Physical tests are commonly used in patellar tendinopathy to aid the clinical diagnosis, assess the prognosis, and monitor treatment. However, it is still unknown whether these physical measures are associated with patient-reported outcomes after exercise therapy.

Purpose: To identify the prognostic value of baseline physical test results and to determine the association between physical response after exercise therapy and clinical improvement over 24 weeks.

Study Design: Case-control study; Level of evidence, 3.

Methods: This study recruited 76 consecutive athletes with patellar tendinopathy who were randomized to 2 different programs of exercise therapy for 24 weeks. Athletes underwent a range of physical tests before and during exercise therapy (12 and 24 weeks), including isometric muscle strength (quadriceps and hip abductors), muscle flexibility (quadriceps, hamstrings, soleus, and gastrocnemius), vertical jump height, and visual analog scale (VAS) scores by palpation, after 3 jump trials, and after single-leg squat (VAS-SLS). The Victorian Institute of Sports Assessment–Patella (VISA-P) questionnaire was used as the primary patient-reported outcome. Linear mixed-effect models were used to assess the prognostic value of baseline physical tests. The change in VISA-P score was further dichotomized into clinical responsiveness (≥14 points) and nonresponsiveness (<14 points). Multiple linear and logistic regression models were performed to evaluate associations between physical response and clinical improvement.

Results: Of the 76 included patients, 67 (88%) had complete follow-ups. The estimated mean VISA-P score increased by 23 points (95% CI, 19-28 points) after 24 weeks. No association was found between any baseline physical test results and a 24-week change in VISA-P score (all Pinteraction > .2, using the likelihood ratio test). Improvement in VAS-SLS after exercise therapy was not associated with VISA-P improvement after adjustment (ß = −1.76; P = .01; Bonferroni-corrected P = .10; R2 = 36.3%). No associations were found between changes in other physical test results and clinical improvement (all P > .05).

Conclusions: In patients with patellar tendinopathy, physical test results including strength and flexibility in the lower limb, jump performance, and pain levels during pain-provoking tests were not identified as prognostic factors for patient-reported outcomes after exercise therapy. Similarly, changes in physical test results were not associated with changes in patient-reported outcomes after adjustments. These results do not support using physical test results to estimate prognosis or monitor treatment response.

Introduction

Patellar tendinopathy (PT) has a prevalence of 17% in the general population[3], and in sports involving jumping, the prevalence rises to 45% in volleyball players and 32% in basketball players[17]. It has a considerable impact on work productivity (impaired in 58% of people with PT) and sports performance (55% of athletes with PT have to reduce or stop their preferred sports activity)[11]. The diagnosis of PT is established when focal pain at the inferior of the patellar is found, especially during tendon-loading activities[19,25]. Exercise therapy programs have been recognized as the most effective approach in conservative treatment for PT[18,25,38].

Physical tests can be used for the functional assessment of PT. These tests, such as muscle strength, flexibility, jump performance, and pain scale during provocation activities, can measure various physical characteristics[10,36]. A range of abnormal physical test characteristics was found to be a risk factor for the onset of PT[24,36,37]. These physical tests are also frequently used in the clinical setting to provide prognosis or to monitor symptoms and function over time in patients with PT. There are, however, currently no studies that evaluated whether physical tests can be used to estimate the course of symptoms (prognosis). Information about changes in the results of physical tests during exercise therapy and its association with the course of symptoms is also lacking. Therefore, the primary aim of this study is to determine the prognostic value of the physical test results measured before starting exercise therapy (baseline) on the change in patient-reported symptoms (difference between baseline and 24 weeks' follow-up). The secondary aim was to assess the association between the change in physical test results and the change in patient-reported symptoms over 24 weeks.

Methods

Study Design and Setting

The study was designed (clinicaltrials.gov, NCT02938143) and conducted at the Erasmus Medical Center (Rotterdam, the Netherlands). The study was approved by the ethics committee of the Erasmus MC University Medical Center, and all patients provided written informed consent before participation. Patients were enrolled from January 2017 to July 2019. For this specific study, we adhered to the Strengthening the Reporting of Observational studies in Epidemiology guidelines for reporting cohort studies and to the minimum reporting standards for tendinopathy studies according to the ICON consensus[31]. This study is a secondary analysis using data from the JUMPER study, a large clinical randomized controlled trial designed to investigate the effect of PTLE (progressive tendon-loading exercises) and EET (eccentric exercise therapy) in athletes with PT for 24 weeks. The PLTE program consisted of 4 stages: isometric, isotonic, energy-storage, and sport-specific exercises. The EET program served as a control group and consisted of 2 stages: a pain-provoking single-leg squat on a decline board with a 25° slope and sport-specific exercises. When patients completed the sport-specific exercises, they were advised to perform maintenance exercises during the return to sports phase. Both exercise programs had a minimum duration of 24 weeks. Additionally, patients were asked to perform exercises targeting risk factors for PT in both study arms. The common treatment principles for tendinopathy involving education and load management advice were given in both groups. Details about the exercise programs, education, and load management advice can be found in the supplementary files of the prior published work[6].

Participants

Consecutive patients aged 18 to 35 years with the diagnosis of PT were enrolled. The main inclusion criteria were a clinical diagnosis of PT (pain on loading and palpation pain on the inferior part of the patella), which had to be confirmed by ultrasound findings (tendon structural changes or anterior-posterior (AP) thickness > 6 mm in grey scale ultrasound and or increased vascularity by power Doppler), and patients with < 80 out of 100 points on the Victorian Institute of Sports Assessment-Patella (VISA-P) questionnaire. The main exclusion criteria were other coexisting knee pathologies and a history of joint injection therapy in the preceding 12 months. Further details about the eligibility criteria are provided in Appendix 1 (available in the online version of this article).

Patient-Reported outcome

The primary outcome was to assess symptom severity change in both groups by a patient-reported outcome measure (PROM) using a translated and validated Dutch VISA-P questionnaire, ranging from 0 to 100 points. A higher score represents fewer symptoms. The minimal clinically important difference (MCID) for the change of VISA-P score among athletes with PT is 14 points or higher[12]. The VISA-P questionnaire was completed by patients at baseline and 12 and 24 weeks follow-up.

Assessment of Physical Tests

Physical test outcomes were collected successively on the same day after the patients completed the VISA-P questionnaire. The coordinating investigator (S.J.B.) assessed the physical test results at baseline and 12 and 24 weeks. A description of the specific physical tests is provided below and shown in Figure 1.

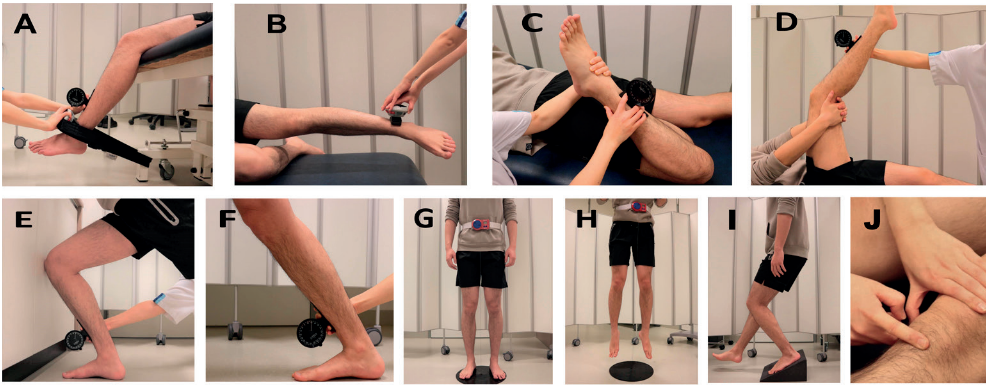


Figure 1. (A) Maximal isometric voluntary contraction (MVC) of the quadriceps muscles was measured with the participant in a seated position on the examination table with both legs hanging over the edge of the table. A fixation band and a plurimeter were used to perform the measurements. Participants were asked to have a straight back and to put both hands on their shoulders. Once the knee flexion angle was at 60° (measured by a plurimeter), the fixation belt was fixed at this position. Participants were asked to put pressure on the dynamometer by contracting their quadriceps while the examiner stabilized the dynamometer. (B) MVC of the hip abductor muscles was measured with the participant in a side-lying position, with the contralateral leg at 90° of flexion. The participant was asked to support his or her head with one hand, and the other hand was used to grasp the edge of the examination table. The participants was asked to put pressure on the dynamometer by contracting the hip abductor muscles against the resistance of the examiner. (C) Quadriceps flexibility was conducted in a prone position. The examiner identified the maximum passive knee flexion angle. Anterior pelvic tilt motion was prevented. (D) Hamstring flexibility was performed in a supine position. The maximum active knee extension angle was dictated by the patient, with a fully extended contralateral leg. (E and F) Participants were asked to lunge in front of the wall. The maximum ankle dorsiflexion range of motion was measured. Soleus flexibility was measured when participants were asked to lunge the knee and touch the wall as far as possible without lifting the heel. The gastrocnemius flexibility was measured by straightening the knee with the ankle at the maximum dorsiflexion angle, still without lifting the heel. (G and H) Participants were asked to wear an adjusted belt connected by a piece of rope before performing 3 trials of the vertical jump. Then, they were asked to land on the mat and report their visual analog scale (VAS) score for maximum pain afterward. The height (cm) was shown on the screen of the device. (I) The single-leg squat was conducted on a 25° decline board. Participants were asked to keep an upright trunk and squat up to 90° of knee flexion. Their VAS score for maximum pain was reported afterward. (J) Tenderness by palpation was assessed using patellar tilting and palpation with the other hand.

Lower Limb Muscle Strength. The maximal isometric voluntary contraction (MVC) was used to assess the strength of the quadriceps and hip abductor muscles using a hand-held dynamometer (MicroFet 2; Hoggan Health Industries)[30]. The isometric contraction was held for 3 seconds and delivered against the resistance by the examiner. The highest score over 2 MVC trials was recorded (N) and normalized by body weight (kg) for analysis. The strength of the quadriceps muscles was assessed with the patients seated and the knee at 60° flexion (Figure 1A). In addition, hip abductor muscle strength was measured in the side-lying position (Figure 1B). Both tests have shown to have good intra-observer reliability (intraclass correlation coefficient [ICC] ranging from 0.73 to 0.98)[5,28].

Lower Limb Flexibility. The flexibility of the lower limb was measured using a plurimeter (Dr. Rippstein, La Conversion). The flexibility of the quadriceps muscles was assessed with patients in a prone position by measuring the maximum passive knee flexion angle[33] (Figure 1C). The reliability of this test in PT is still unknown. The active knee extension test was used to measure the flexibility of the hamstring muscles with patients in a supine position (Figure 1D). This test has been reported to have good intra-observer reliability (ICC 0.84) in PT[10]. The flexibility of the soleus muscle (or ankle dorsiflexion angle) was measured by assessing the angle of the tibial shaft during a weight-bearing dorsiflexion lunge test (Figure 1E). This test has shown to be excellent intra-observer reliability (ICC 0.98)[10]. The flexibility of gastrocnemius (Figure 1F) was evaluated by calculating the maximum ankle dorsiflexion in the weight-bearing technique with the knee extended. This test has shown to have good intra-observer reliability (ICC from 0.77 to 0.89) and high inter-observer reliability (ICC 0.95)[27].

Performance Test. The maximum jump height was used as a performance test (Figure 1, G and H). We did not instruct a specific technique. The patients had to jump as high as possible in a vertical direction using both legs (a countermovement prior to the jump was allowed). The maximal jump height out of 3 trials[23] was recorded in centimeters using a digital vertical jump meter (Takei 5406 Jump-MD, Takei Scientific Instruments Co). The intra-observer reliability of vertical jumps ranges from 0.93 to 0.98[23].

Pain Provocation Tests. A Visual Analogue Scale (VAS) score from 0 to 10 points was used to quantify pain levels during specific tests (Figure 1,H-J). Pain provocation tests included VAS on palpation (VAS-palpation), after three jump trials (VAS-3-jumps), and single-leg squat (VAS-SLS)[29]. The order of these tests was always the same, with VAS-palpation being the first pain provocation test and VAS-SLS being the last pain provocation test. The maximal VAS score during or directly after these tests was recorded. The reliability of these tests in PT is unknown.

Statistical Analysis

The first author (J.D.) was conducting statistical analyses and was not involved in performing the physical tests or collecting the VISA-P scores. Statistical analyses under intention-to-treat were performed. Normality tests were performed by visual check using histograms or normal Q-Q plots and also assessed by the Shapiro-Wilk's test.

Estimated means and standard deviations were reported for continuous variables with normal distribution. Otherwise, medians and interquartile ranges (IQRs) were used. We performed a two-sample t-test and Mann-Whitney *U* test to determine between-group differences at baseline as appropriate.

We use a linear mixed-effect model to evaluate the change in clinical symptoms (VISA-P score) and physical test results in both groups over 24 weeks. The fixed-effect included time, study arms, baseline clinical characteristics (symptom duration, sports activity [Cincinnati Sports Activity Scale (CSAS)][4], sex, BMI, and age. Individual participant level and time (constant up to 24 weeks) were included as a random effect. In addition to these basic models, each baseline physical test result was added to investigate their association with the estimated change in VISA-P score. Interaction between each of these physical test results and time was tested and retained in the model only if the p-value was significant (<.05) by likelihood ratio test (LRT). To visualize these associations for illustration purposes, we used an arbitrary split using the median value in each of the baseline physical test results. No sensitivity analysis was performed as the LMM can provide an unbiased result under the assumption that missing data occurred at random.

Linear regression models were used to assess the association between the change in physical test results and the change in VISA-P score over 24 weeks. In addition, we further dichotomized the change in VISA-P score into clinical responsiveness (≥14 points), and non-responsiveness (< 14 points) by MCID, and logistic regression analyses were further used to assess whether these physical changes were associated with the occurrence of clinical responsiveness. All models were adjusted with aforementioned baseline clinical predictors, and additional baseline VISA-P score. In case of missing data in VISA-P scores or physical test results of >5%, prediction linear mixed models were utilized to impute missing values (Appendix A1, available online). We also performed a sensitivity analysis based on a complete case dataset, in which missing data were excluded.

Residuals were tested to investigate the assumptions of models. Collinearity and outliers were also tested for the linear regression model. In case of multiple testing, a Bonferroni correction method was used[13]. P values (<.05) were considered statistical significant. All analyses were done by R software, version (2022.4.2.1) (The R Foundation for Statistical Computing). Used packages included “nlme”, ”glm”, ”emmeans”, “car”, ”ggplot2”, and ”psych”.

Results

Patient Characteristics

A total of 76 athletes were included. A summary of baseline characteristics is shown in Table 1, using a recommended report format in tendinopathy research[31]. Nine athletes (12%) were lost to follow-up at 24 weeks (Figure 2). For baseline physical tests, there were missing values only for VAS-palpation (n=11 missing; 14%). In addition, baseline physical test results and VISA-P score between clinical responsiveness and non-responsiveness after 24 weeks are also presented in Appendix A2 (available online). No imbalances were found in physical test results at baseline.

Table 1. Baseline Characteristics of 76 Athletes[a]

	Value
Demographics	
Age, y, mean (SD)	25 (4)
Male sex	58 (76)
Weight, kg, mean (SD)	81.8 (12.3)
Height, cm, mean (SD)	184.7 (9.3)
BMI, kg/m^2, mean (SD)	23.9 (2.9)
Tendinopathy descriptors	
Bilateral tendon involvement	32 (42)
Symptom duration, wk, median [IQR]	104 [49-208]
VISA-P score, mean (SD)	55 (13)
CSAS sports activity level[b]	
Level 1: 4-7 d/wk)	17 (22)
Level 2: 1-3 d/wk)	59 (78)
Imaging for assisting diagnosis	
Use of ultrasound and MRI	76 (100)
Medication use	
Use of painkillers	11 (15)
General Health	
EQ VAS, median [IQR][c]	85 [75 – 90]

[a]Data are presented as n (%) unless otherwise indicated. BMI, body mass index; CSAS, Cincinnati Sports Activity Scale; EQ, Euro Qol; MRI, magnetic resonance imaging; VAS, visual analog scale; VISA-P, Victorian Institute of Sports Assessment–Patella.

[b]According to Barber-Westin et al.[4]

[c]The EQ VAS is part of the EQ-5D and measures a patient's perception of one's overall current health on a vertical VAS (0-100 points), with a higher score representing better general health status.

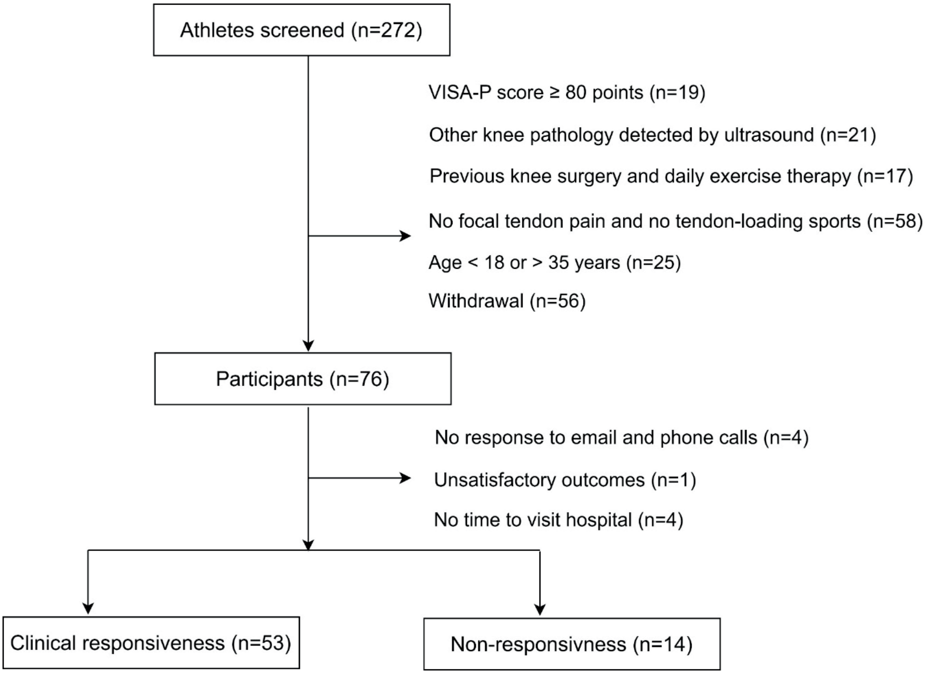


Figure 2. Flow diagram. VISA-P, Victorian Institute of Sports Assessment–Patella.

Prognostic Value of Baseline Physical Results

There was a significant increase in VISA-P score in the whole group, with an increase of 23 points (95% CI: 19 to 28 points; Bonferroni-corrected $P < .001$) from baseline to 24 weeks (Appendix A3, available online). To evaluate whether any baseline physical test result is associated with the increase in VISA-P score over 24 weeks, we tested interactions of time with each of the physical results in each multivariable linear mixed-effect model (Table 2). The results showed that these baseline physical test results were not significant prognostic factors of VISA-P improvement (all $P_{interaction} > 0.2$), despite lower VAS-3-jumps scores being significantly associated with higher VISA-P scores at 12 or 24 weeks (ß = -2.05; 95% CI, -3.06 to -1.04; Bonferroni-corrected $P < .01$). These results were also visualized by dichotomizing baseline physical test results into low and high levels using the median value for illustration purposes, indicating the rate of increase in VISA-P score throughout the 24 weeks was constant in all subgroups according to baseline physical test results level (Figure 3 and Appendix 4, available online).

Table 2. Prognostic mixed-effect model of baseline physical test results for VISA-P improvement[a]

Tests	Time X physical test resultsb		physical test resultsc	
	P valued	LRT	ß (95% CI)	*P* valued
Quadriceps strength, N/kg	.258	1.28	1.39 (-1.33 to 4.12)	.322
Hip abductors strength, N/kg	.327	0.96	1.00 (-5.42 to 7.42)	.761
Quadriceps flexibility, degrees	.323	0.98	-0.18 (-0.59 to 0.23)	.393
Hamstrings flexibility, degrees	.852	0.03	-0.10 (-0.36 to 0.16)	.461
Soleus flexibility, degrees	.417	0.66	0.10 (-0.30 to 0.50)	.623
Gastrocnemius flexibility, degrees	.395	0.72	-0.19 (-0.57 to 0.18)	.321
Vertical jump height, cm	.675	0.18	0.12 (-0.22 to 0.46)	.505
VAS-palpation	.834	0.04	-0.38 (-1.48 to 0.72)	.507
VAS-3-jumps	.750	0.10	-2.05 (-3.06 to -1.04)	<.001
VAS-SLS	.867	0.03	-1.64 (-2.81 to -0.47)	.008

[a] VISA-P, Victorian Institute of Sports Assessment-patella(0-100 point); VAS, Visual Analog Scale (rating scale [0-10]); VAS-palpation, VAS by palpation test; VAS-3-jumps, VAS after 3 three jump trials; VAS-SLS, VAS after single leg squat test; LRT, likelihood ratio statistic.

[b]Interactions terms (time x each physical test result) was tested using likelihood ratio test, representing the effect of physical test results on VISA-P score over time.

[c]The effect of physical test results from models without the time x physical tests interaction.

[d]These p values were not adjusted by multiple testing. The Bonferroni corrected p-value is using multiplying the raw p values by the number of the tests (n=10). P values < .05 were considered statistically significant.

Association Between the Changes in Physical Test Results and Change in VISA-P Score

The longitudinal changes in each physical test result are summarized in Appendix 3 (available online). The aforementioned physical tests, including muscle strength, flexibility, and pain provocation results, improved significantly over time (all Bonferroni-corrected $P < .001$), while no statistically significant longitudinal change in jump height was found (ß = 1; 95% CI, –1 to 2; Bonferroni-corrected*P* = .836).

In the multivariable linear regression models, only a decreased VAS-SLS score was significantly associated with an increase in VISA-P score (ß = –1.76; 95% CI, –3.09 to –0.43; P = .01; $R2$ = 36.3%) (Figure 4A). However, this association did not reach statistical significance after Bonferroni correction (P = .1). No associations were found between other physical test result changes and VISA-P improvement (Appendix 5, available online). Furthermore, 58 athletes (76%) were categorized as clinically responsive based on the VISA-P score criterion (change of VISA-P score ≥14 points after 24 weeks), and 18 athletes (24%) as nonresponsive.

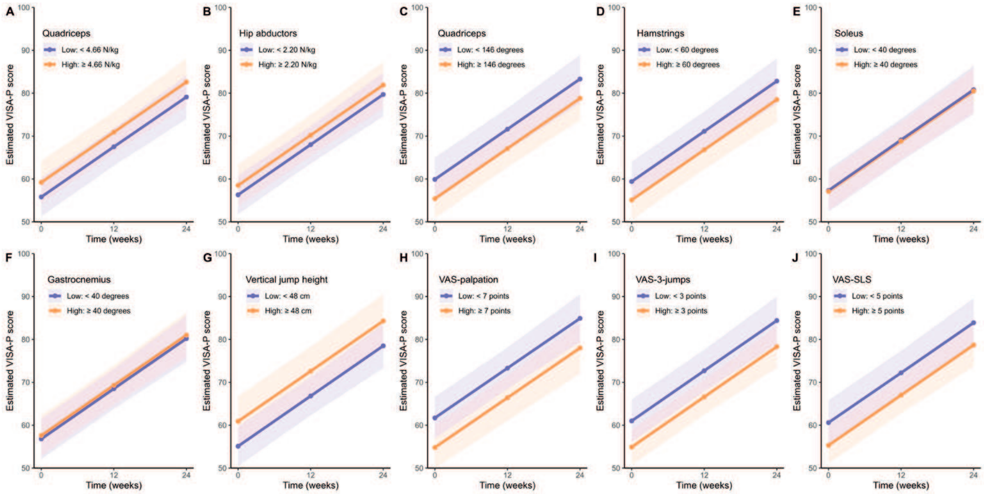


Figure 3. (A-J) Visualization of the prognostic value of baseline physical test results.

The prognostic value of each baseline physical test result on the progression of the Victorian Institute of Sports Assessment–Patella (VISA-P) score (0-100 points) over 24 weeks is depicted by dichotomizing baseline physical test results into low and high levels using the median value for illustration purposes. Plots represent the progression of VISA-P scores over 24 weeks by different levels of baseline physical test results. Dots represent marginal estimated means by mixed-effect model without time × physical test interaction. Shaded areas represent the 95% CIs after the Bonferroni correction. There was not a significant difference in the rate of increase in VISA-P between low and high baseline physical test results. VAS, visual analog scale; VAS-palpation, VAS by palpation test; VAS-SLS, VAS after single-leg squat test; VAS-3-jumps, VAS after 3 jump trials.

We ran supplementary analyses on the association between the change in physical test results and the occurrence of clinical responsiveness. In the multivariable logistic regression models, no significant associations were found between the change in any physical test result and clinical responsiveness (all $P > .05$) (Figure 4). Sensitivity analyses performed by complete-case data confirmed the robustness of our outcomes (Appendix 5, available online).

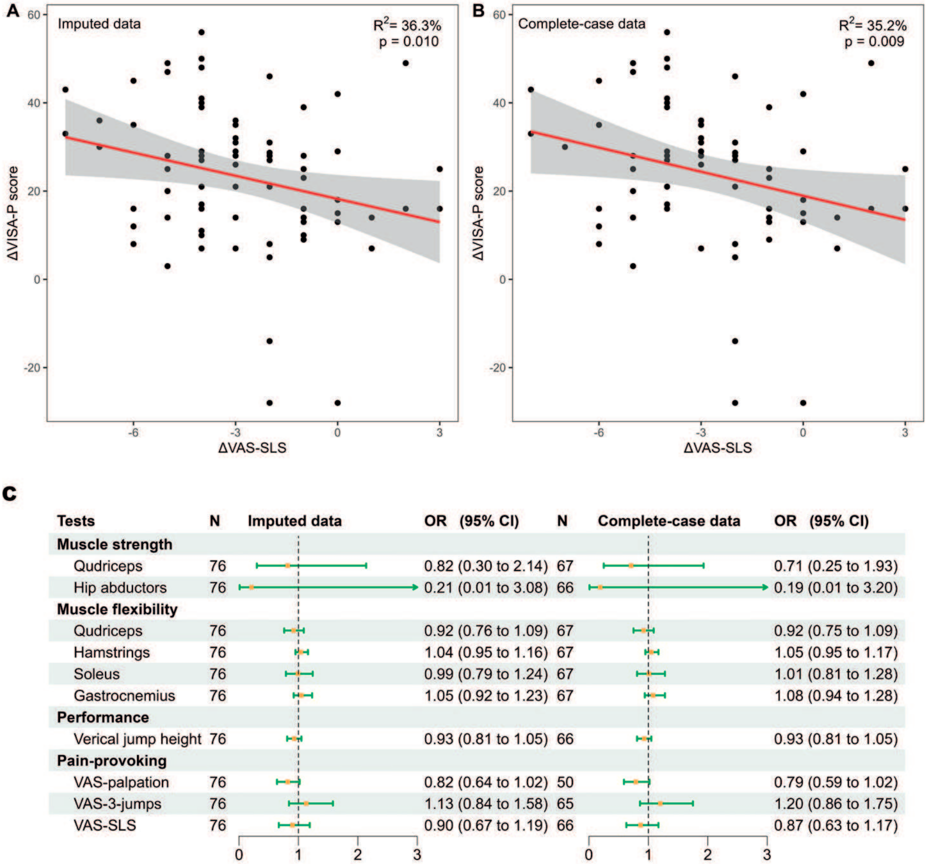


Figure 4. The association between change in physical test results and clinical improvement. (A and B) The associations between the change in the Victorian Institute of Sports Assessment–Patella (ΔVISA-P) score and the change in the visual analog scale single-leg squat test (ΔVAS-SLS) score were evaluated by linear regression models using imputed and complete-case data, adjusted with study arms, sex, body mass index, age, and previous symptom duration, sports activity, and baseline VISA-P scores. Lines represent regression lines. Shaded areas represent the 95% CIs. The P values are not adjusted by Bonferroni correction. (C) Forest plots of multivariable logistic regression to make an association between the change in physical test results and the occurrence of clinical responsiveness (change of VISA-P score ≥14 points) using imputed and complete-case data sets. OR, odds ratio; VAS-palpation, VAS by palpation test; VAS-3-jumps, VAS after 3 jump trials.VAS after single leg squat test; R^2, adjusted R-squared by linear regression models; Δ, the change from baseline to 24 weeks; OR, odd ratio.

Discussion

This is the first large-scale study unraveling the association between physical test results and change in a patient-reported outcome in athletes with PT. We found that none of the baseline physical test results were identified as a prognostic factor on clinical outcome after exercise treatment. Specifically, baseline differences in lower limb muscle strength, flexibility, jump performance, and pain level during pain-provoking tests were not associated with an improvement in VISA-P score over 24 weeks. In addition, for athletes receiving exercise therapy over time, we observed a significant but small increase in isometric muscle strength of quadriceps and hip abductors, lower limb flexibility, and decreased pain level during provocation tests. Furthermore, a decreased VAS-SLS was associated with an increase in VISA-P, while there was a lack of statistical significance after multiple testing corrections. Other physical test changes after exercise treatment were not associated with the change in patient-reported outcomes after 24 weeks.

These findings are essential because prognostic factors for predicting clinical outcomes following exercise treatment for PT are lacking. In previous studies, some physical test results have been identified as risk factors for developing PT. These include weakness in quadriceps and hip abductor strength, reduced quadriceps, hamstrings and lower ankle flexibility, and greater vertical jump performance [8,24,36,37]. While these data could be valuable to inform studies on prognostic factors, it should be stressed that risk factors and prognostic factors are different entities. In the clinical setting, pain provocation tests, such as palpation or single-leg squat, are widely used to aid in confirming a diagnosis. It is therefore relevant to investigate whether these easy-to-perform tests could aid in providing a prognosis or monitoring the condition. To date, exercise treatment has been suggested as the first-line therapy for symptomatic PT[21,25]. However, subgroups of patients will not benefit from exercise therapy[15,18]. Detecting possible prognostic factors could not only help clinicians to estimate the prognosis but also help develop the decision rule to select subgroups of patients with those characteristics who would most likely benefit from exercise treatment for PT[14].

Prognostic Value of Baseline Physical Test Results

To our knowledge, this is the first study to evaluate whether pretreatment physical test results can predict a 24-week change in VISA-P score for athletes with PT performing exercise therapy. A previous cohort study had identified a set of demographic predictors of 14-week VISA-P change following exercise treatment for PT[38]. The prognosis was worse when participants with older age, a longer duration of symptoms, and higher training volume following eccentric exercise. In our study, after adjusting these predefined demographic predictors and exercise types, we found no association between these pretreatment physical tests and a 24-week change in VISA-P score. Furthermore, we did not detect any difference in baseline physical test results between athletes who achieve clinical responsiveness (the minimal change in VISA-P score ≥ 14 points) and non-responsiveness (the minimal change in VISA-P score < 14 points). Collectively, these findings suggest that these pretreatment physical test results may not influence the course of clinical response to exercise treatment. A possible explanation for this finding is that, rather than physical features, other stronger prognostic factors (such as psychological factors) might be crucial components for treatment effectiveness in tendinopathy[20]. Because exercise therapy can induce pain, it might result in fear and prevent achieving meaningful clinical changes in some patients[16,25]. One cross-sectional study reported psychological outcomes, such as pain catastrophizing and kinesiophobia, had a negative influence on the VISA score[34]. Thus, it is reasonable to assume that even if initial physical impairments may affect clinical outcomes, their effect is probably counteracted by these more essential prognostic factors.

Association Between Physical Response and Clinical Improvement

Clinical improvement following exercise therapy has been shown both in the short-term and long-term follow-up[1,6,32]. However, whether these clinical improvement is a results of the improvement in physical properties is still unclear. The VISA-P score is the primary or most commonly used outcome to quantify treatment response in patients with PT[7], incorporating pain, function and ability to play sports. Knowing this relationship will provide information on whether these physical tests could be used to monitor the effectiveness of exercise therapy.

Lower Limb Muscle Strength and Flexibility. We observed a constant but small increase in lower limb muscle strength and flexibility following exercise therapy over 24 weeks. Several studies have also reported physical change only with a shorter follow-up duration (12 weeks)[1,32,35]. Agergaard el al[1] found an increased quadriceps strength for PT following exercise therapy, opposite to Sprague et al[35] and Ruffino et al[32], who reported no change in this muscle strength. In addition, only 1 study[32] reported muscle flexibility (ankle dorsiflexion) following exercise therapy, and no change was identified. These discrepancies in muscle response may be due to the difference exercise programs used in the abovementioned studies. However, we found that these physical improvements were not associated with VISA-P increase or meaningful clinical change. A potential explanation for the lack of association may be the small magnitude of change value observed in physical tests, indicating that more sensitive and advanced equipment may be needed to detect changes in muscle strength response during exercise treatment.

Performance Test. As for vetitical jump height, despite the fact that Sprague et al[35] reported a jump height improvement after 12 weeks, our data showed no change was observed in vertical jump height, similar to other previous studies[1,32]. It is interesting to note that jump performance in our study population did not improve given that muscle strength and flexibility were both increased following exercise therapy; that may be because other factors, such as arm swing or trunk angle, can also affect vertical jump height[22]. Also, psychological readiness might affect performing a maximum jump in athletes recovering from PT[2].

Pain Provocation Tests. In line with a previous study[1], we also found a reduced VAS score during pain-provoking tests over 24 weeks in our study population. However, only the decline in VAS-SLS was associated with the change in VISA-P score over 24 weeks. As most of the VISA-P items comprise pain scores, it may be reasonable to assume that the VAS-SLS decline may predict clinical improvement. However, it is important to note that this result was not statistically significant after the Bonferroni correction, in the setting of a relatively moderate strength of association between these two outcomes (R^2 = 36.3%). One possible explanation is that while pain reduction is a critical goal of treatment for PT, it may not necessarily related to a reduction in the underlying tendon pathology which is more directly reflected by the VISA-P score[9]. Furthermore, none of the pain reduction during these tests was associated with the probability of achieving meaningful clinical changes. Taken together, our outcomes suggest that pain provocation test results may not be useful for solely monitoring treatment response.

5

Strengths and Limitations

The major strength of our study is that this is a large prospective study performed in athletes with PT, with structured data collection by a single trained researcher (S.J.B.). The VISA-P score was administered in a standardized way before performing physical tests. Thus, a potential influence on physical test results was avoided as much as possible. Most of the physical test methods have good reliability, and measurement was conducted by the same person, which can reduce the measurement error. Additionally, we performed a sensitivity analysis based on the imputation of data and complete-case data, which enhanced the robustness of our results.

The limitations of our study should also be addressed. First, although a prospective study is the preferable design to answer prognostic questions, a randomized trial can be implemented by combining intervention and control group and then adjusting the treatment variable in the prognostic model[26]. However, the generalizability may be reduced when these results are used in athletes whose characteristics are not similar to those of the enrolled participants, because strict inclusion criteria are used in a trial. Second, in our study, other exercise regimes such as isolated heavy slow resistance and isometric exercises[18] were not included. We strongly believe this does not limit the main findings because there is no strong evidence that one exercise therapy is more beneficial than others[6,21,25]. Third, we adjusted *P* values using the Bonferroni correction in multiple testing, which may be too conservative to reject true associations.

Conclusion

In this large prospective cohort study, we identified that none of the baseline physical test results were associated with the change in patient-reported outcome over 24 weeks in athletes with PT after exercise treatment. Although muscle strength, flexibility, and pain level were improved during exercise treatment, no association was found between the change in these physical test results and symptom improvement. Only an improvement in the VAS-SLS test was associated with an improvement in the change in patient-reported outcome, although this association was not statistically significant after adjustment. These results do not support using physical test results to estimate prognosis or monitor treatment response. These findings aid physicians in making an adequate interpretation of the value of physical test results during the recovery of athletes with PT.

References

1. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, et al. Clinical Outcomes, Structure, and Function Improve With Both Heavy and Moderate Loads in the Treatment of Patellar Tendinopathy: A Randomized Clinical Trial. *Am J Sports Med*. 2021;49(4):982-993. PMID: 33616456
2. Aizawa J, Hirohata K, Ohji S, Ohmi T, Koga H, Yagishita K. Factors Associated With Psychological Readiness to Return to Sports With Cutting, Pivoting, and Jump-Landings After Primary ACL Reconstruction. *Orthop J Sports Med*. 2020;8(11):2325967120964484. PMID: 33244476
3. Albers IS, Zwerver J, Diercks RL, Dekker JH, Van den Akker-Scheek I. Incidence and prevalence of lower extremity tendinopathy in a Dutch general practice population: a cross sectional study. *BMC Musculoskelet Disord*. 2016;17:16. PMID: 26759254
4. Barber-Westin SD, Noyes FR. Assessment of sports participation levels following knee injuries. *Sports Med*. 1999;28(1). PMID: 10461708
5. Boling MC, Padua DA, Marshall SW, Guskiewicz K, Pyne S, Beutler A. A prospective investigation of biomechanical risk factors for patellofemoral pain syndrome: the Joint Undertaking to Monitor and Prevent ACL Injury (JUMP-ACL) cohort. *Am J Sports Med*. 2009;37(11):2108-2116. PMID: 19797162
6. Breda SJ, Oei EHG, Zwerver J, et al. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. *Br J Sports Med*. 2021;55(9):501-509. PMID: 33219115
7. Chen P-C, Wu K-T, Chou W-Y, et al. Comparative Effectiveness of Different Nonsurgical Treatments for Patellar Tendinopathy: A Systematic Review and Network Meta-analysis. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association*. 2019;35(11). PMID: 31699265
8. Cook JL, Kiss ZS, Khan KM, Purdam CR, Webster KE. Anthropometry, physical performance, and ultrasound patellar tendon abnormality in elite junior basketball players: a cross-sectional study. *Br J Sports Med*. 2004;38(2):206-209. PMID: 15039260
9. Cook JL, Rio E, Purdam CR, Docking SI. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? *Br J Sports Med*. 2016;50(19):1187-1191. PMID: 27127294
10. Crossley KM, Thancanamootoo K, Metcalf BR, Cook JL, Purdam CR, Warden SJ. Clinical features of patellar tendinopathy and their implications for rehabilitation. *J Orthop Res*. 2007;25(9):1164-1175. PMID: 17469196
11. De Vries AJ, Koolhaas W, Zwerver J, et al. The impact of patellar tendinopathy on sports and work performance in active athletes. *Res Sports Med*. 2017;25(3):253-265. PMID: 28391723
12. Hernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *Br J Sports Med*. Mar 2014;48(6):453-7. PMID: 23012320
13. Jafari M, Ansari-Pour N. Why, When and How to Adjust Your P Values? *Cell J*. 2019;20(4):604-607. PMID: 30124010
14. Kent P, Cancelliere C, Boyle E, Cassidy JD, Kongsted A. A conceptual framework for prognostic research. *BMC Med Res Methodol*. 2020;20(1):172. PMID: 32600262
15. Kettunen JA, Kvist M, Alanen E, Kujala UM. Long-term prognosis for jumper's knee in male athletes. A prospective follow-up study. *Am J Sports Med*. 2002;30(5):689-692. PMID: 12239003
16. Kvist J, Silbernagel KG. Fear of Movement and Reinjury in Sports Medicine: Relevance for Rehabilitation and Return to Sport. *Phys Ther*. 2022;102(2). PMID: 34971375
17. Lian OB, Engebretsen L, Bahr R. Prevalence of jumper's knee among elite athletes from different sports: a cross-sectional study. *Am J Sports Med*. 2005;33(4):561-567. PMID: 15722279

18. Lim HY, Wong SH. Effects of isometric, eccentric, or heavy slow resistance exercises on pain and function in individuals with patellar tendinopathy: A systematic review. *Physiother Res Int*. 2018;23(4):e1721. PMID: 29972281
19. Malliaras P, Cook J, Purdam C, Rio E. Patellar Tendinopathy: Clinical Diagnosis, Load Management, and Advice for Challenging Case Presentations. *J Orthop Sports Phys Ther*. 2015;45(11):887-898. PMID: 26390269
20. Mallows A, Debenham J, Walker T, Littlewood C. Association of psychological variables and outcome in tendinopathy: a systematic review. *Br J Sports Med*. 2017;51(9):743-748. PMID: 27852585
21. Marigi EM, Buckley P, Razi F, et al. Patellar Tendinopathy: Critical Analysis Review of Current Nonoperative Treatments. *JBJS Rev*. 2022;10(3). PMID: 35358114
22. Markovic G. Does plyometric training improve vertical jump height? A meta-analytical review. *Br J Sports Med*. 2007;41(6). PMID: 17347316
23. Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. *J Strength Cond Res*. 2004;18(3):551-555. PMID: 15320660
24. Mendonça LD, Ocarino JM, Bittencourt NFN, Macedo LG, Fonseca ST. Association of Hip and Foot Factors With Patellar Tendinopathy (Jumper's Knee) in Athletes. *J Orthop Sports Phys Ther*. 2018;48(9):676-684. PMID: 29792104
25. Millar NL, Silbernagel KG, Thorborg K, et al. Tendinopathy. *Nat Rev Dis Primers*. 2021;7(1):1. PMID: 33414454
26. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ (Clinical Research ed)*. 2009;338:b375. PMID: 19237405
27. Munteanu SE, Strawhorn AB, Landorf KB, Bird AR, Murley GS. A weightbearing technique for the measurement of ankle joint dorsiflexion with the knee extended is reliable. *J Sci Med Sport*. 2009;12(1):54-59. PMID: 17888733
28. Piva SR, Fitzgerald K, Irrgang JJ, et al. Reliability of measures of impairments associated with patellofemoral pain syndrome. *BMC Musculoskelet Disord*. 2006;7:33. PMID: 16579850
29. Purdam CR, Cook JL, Hopper DM, Khan KM, group VISts. Discriminative ability of functional loading tests for adolescent jumper's knee. *Physical Therapy in Sport*. 2003/02/01/ 2003;4(1):3-9.
30. Rathleff MS, Samani A, Olesen JL, Roos EM, Rasmussen S, Madeleine P. Effect of exercise therapy on neuromuscular activity and knee strength in female adolescents with patellofemoral pain-An ancillary analysis of a cluster randomized trial. *Clin Biomech (Bristol, Avon)*. 2016;34:22-29. PMID: 27054583
31. Rio EK, Mc Auliffe S, Kuipers I, et al. ICON PART-T 2019-International Scientific Tendinopathy Symposium Consensus: recommended standards for reporting participant characteristics in tendinopathy research (PART-T). *Br J Sports Med*. 2020;54(11):627-630. PMID: 31519545
32. Ruffino D, Malliaras P, Marchegiani S, Campana V. Inertial flywheel vs heavy slow resistance training among athletes with patellar tendinopathy: A randomised trial. *Phys Ther Sport*. 2021;52:30-37. PMID: 34384941
33. Sanchis-Alfonso V, Coloma-Saiz J, Herrero-Herrero M, Prades-Piñón J, Ramírez-Fuentes C. Evaluation of anterior knee pain patient: clinical and radiological assessment including psychological factors. *Ann Jt*. 2018;3(3).
34. Slagers AJ, van Veen E, Zwerver J, Geertzen JHB, Reininga IHF, van den Akker-Scheek I. Psychological factors during rehabilitation of patients with Achilles or patellar tendinopathy: a cross-sectional study. *Phys Ther Sport*. 2021;50:145-152. PMID: 34015607

35. Sprague AL, Couppé C, Pohlig RT, Snyder-Mackler L, Silbernagel KG. Pain-guided activity modification during treatment for patellar tendinopathy: a feasibility and pilot randomized clinical trial. *Pilot Feasbility Stud.* 2021;7(1):58. PMID: 33632313
36. Sprague AL, Smith AH, Knox P, Pohlig RT, Grävare Silbernagel K. Modifiable risk factors for patellar tendinopathy in athletes: a systematic review and meta-analysis. *Br J Sports Med.* 2018;52(24):1575-1585. PMID: 30054341
37. van der Worp H, van Ark M, Roerink S, Pepping G-J, van den Akker-Scheek I, Zwerver J. Risk factors for patellar tendinopathy: a systematic review of the literature. *Br J Sports Med.* 2011;45(5):446-452. PMID: 21367808
38. van Rijn D, van den Akker-Scheek I, Steunebrink M, Diercks RL, Zwerver J, van der Worp H. Comparison of the Effect of 5 Different Treatment Options for Managing Patellar Tendinopathy: A Secondary Analysis. *Clin J Sport Med.* 2019;29(3):181-187. PMID: 31033610

Supplementary material

Supplementary material for this paper can be found online:

https://doi.org/10.1177/03635465231200241

6

Chapter 6

Prognostic Value of Conventional Ultrasound and MRI Features for Clinical Outcomes in Athletes with Patellar Tendinopathy Following Exercise Therapy

Jie Deng, Stephan J Breda, Yijie Fang, Denise Eygendaal, Robert-Jan de Vos, Edwin HG Oei

Accepted, Sports Health: A Multidisciplinary Approach

Abstract

Background: Structural abnormalities assessed with conventional ultrasound (US) or MRI are associated with the risk of developing patellar tendinopathy (PT). However, their prognostic value for athletes with PT performing exercise therapy remains unclear.

Purpose: To assess the association between baseline imaging features and change in pain and disability over 24 weeks in athletes with PT following exercise treatment.

Study Design: Cohort Study

Methods: This study is a secondary analysis of a randomized trial involving athletes with PT who were randomly allocated to 2 different programs of exercise therapy for 24 weeks. Imaging features at baseline included patellar tendon thickness, intratendinous calcifications, patellar erosions, and Doppler flow on US, as well as tendon fiber disruption, infrapatellar fat pad (IFP) edema, bone marrow edema, and deep infrapatellar bursitis on MRI. Clinical outcomes were measured at baseline, 12- and 24-week follow-up, using the Visual Analog Scale after single-leg squat (VAS-SLDS) for pain on loading, and Victorian Institute of Sports Assessment–Patella (VISA-P) questionnaire for disability. Linear mixed-effects models, incorporating interaction terms tested using likelihood ratio tests (LRT), were used to evaluate the prognostic value of baseline imaging features.

Results: Among the 76 included athletes (58 male, 18 female, average age 24 ± 4 years), abnormal US features were identified in 26-78% of cases. Among the MRI analyzed (n = 72), abnormal MRI features were demonstrated in 43-81% of cases. No significant associations were identified between individual imaging features and 24-week changes in VAS-SLDS or VISA-P scores (all $P_{interaction} > 0.1$), nor between the total number of imaging abnormalities and these clinical outcomes (all $P_{interaction} > 0.5$).

Conclusions: There was no evidence of an association between baseline abnormalities assessed using conventional ultrasound or MRI and 24-week changes in pain or disability among athletes with PT undergoing exercise therapy. Healthcare professionals should avoid relying on these imaging findings to predict prognosis.

Introduction

Patellar tendinopathy (PT) is a highly prevalent injury in jumping athletes[33]. This disease is characterized by long-lasting pain and disability and typically has a negative impact on sports performance and work participation[18]. Exercise therapy is considered the first-line treatment in PT[13,17,31]. The clinical course of pain and disability in PT athletes following this standard treatment is often characterized by modest improvement in average pain and disability at 3 months of follow-up[12], with little or no further change during long-term[1]. Substantial inter-individual variations exist in these outcomes, suggesting that some athletes experience meaningful recovery while others show minimal response[1,19,31]. This variability highlights the need for prognostic research to identify factors[36] that can help distinguish between individuals with a favorable and unfavorable treatment trajectory. This would facilitate patient education and optimize treatment recommendations.

Patients with PT commonly demonstrate structural changes or increased vascularization within the patellar tendon. Previous studies[15,30] have found that tendon thickening and neovascularization on ultrasound (US) are associated with a higher risk of developing PT. Additionally, US-detected intratendinous calcification or patella erosions in PT patients may indicate the level of disease chronicity[10]. Magnetic resonance imaging (MRI) is also used in PT to identify pathological changes in adjacent structures such as infrapatellar fat pad (IFP, or Hoffa's fat pad)[14] edema, and bone marrow edema[34]. Besides, the ratio of partial tear thickness to overall tendon thickness on MRI (i.e., > 50% tear thickness) may serve as a classifier for surgical decision-making in athletes with PT[23]. While these features are commonly used to assist in clinical diagnosis, their prognostic value for predicting clinical outcomes following exercise treatment in PT athletes remains poorly understood.

Therefore, this study aimed to explore the prognostic value of structural abnormalities detected by conventional US and MRI at baseline on the 24-week change in pain and disability in athletes with PT following exercise treatment.

Methods

Study Design and Participants

This is a secondary analysis of a randomized controlled trial[7] of 76 athletes with PT performing progressive tendon-loading exercises (PTLE) or eccentric exercise therapy (EET) at Erasmus University Medical Center between January 2017 and July 2019 (ClinicalTrial.gov: NCT02938143). The PTLE group consisted of isometric loading, isotonic loading, energy-storage loading and sport-specific exercises. Athletes performing the EET program were advised to do a single-leg squat using a 25° decline board and sport-specific exercises. Both groups performed their prescribed exercises for 24 weeks in an unsupervised setting. However, based on their own preference, they arranged supervision by a health professional. Details on the exercise treatment strategies were published previously[7]. All participants provided informed consent for the study, which was approved by the institutional review board (MEC-2016-500). For this secondary analysis, we combined patients undergoing both exercise treatments into one cohort to enhance statistical power and maximize the available data. Given that the intervention group demonstrated a statistically superior clinical improvement at 24 weeks[7], we accounted for potential differences by including treatment allocation as a covariate in the model[32]. We adhered to the minimum reporting standards for tendinopathy trials according to the international consensus statement[38]. This study was performed and reported in accordance with the guidelines for assessing prognostic factors under the PROGRESS framework[36,37].

Consecutive athletes aged 18 to 35 years with PT (characterized by pain on loading and palpation pain on the inferior part of the patella, the Victorian Institute of Sports Assessment-Patella (VISA-P) questionnaire < 80 points) confirmed by US (tendon structural changes and/or anterior-posterior thickness > 6 mm in grayscale ultrasound and/or increased vascularity by power Doppler) were enrolled. Main exclusion criteria included coexisting knee pathologies on ultrasound or MRI (e.g., cartilage lesions, full-thickness patellar tendon rupture), as these could confound the diagnosis or outcome of patellar tendinopathy. We also excluded participants with a history of joint injection therapy or ≥4 weeks of structured exercise therapy in the previous 12 months, to minimize residual effects from prior treatments. Additional exclusion criteria: prior knee surgery without full rehabilitation, previous patellar tendon rupture, acute knee or patellar tendon injuries, or inability to participate in an exercise program, either due to safety concerns or to ensure adherence to the intervention protocol.

Measurements

Clinical outcome measures

Pain and disability related to PT were measured at baseline, 12 weeks, and 24 weeks using pain intensity rated on the Visual Analog Scale directly after the single-leg decline squat (VAS-SLDS) and VISA-P score, respectively. These outcome measures align with a reported consensus for core domains for tendinopathy[44]. Patients performed the single-leg decline squat test when completed the VISA-P questionnaire before the imaging examination was done, so they were not aware of the imaging findings when they reported the clinical severity of their symptoms at that specific time point.

Ultrasound

Grey-scale US (GSUS) and power Doppler US (PDUS) were performed at baseline and conducted in the longitudinal and transverse planes using a linear 5-15MHz transducer (GE Healthcare, ML6-15; LOGIQ E9), with an ultrasound gel (Sonogel Vertriebs GmbH)[6]. All participants were in the supine position with the back rest of the examination table upright in 60° for patient comfort. Both knees were placed in 30° flexion to limit anisotropy, supported by a foam roll. The imaging was acquired with a frequency of 15 MHz, a fixed depth of 2.5 cm, a gain of 41 dB, auto optimization of 100%, and a dynamic range of 66 dB. Patellar tendon and surrounding tissues were evaluated on US as follows[10,21]: 1) anteroposterior (AP) thickness of patellar tendon was measured in the transverse plane at the thickest point within 1 cm distal to the inferior pole of the patella, 2) intratendinous calcifications: hyperechoic within patellar tendon with posterior acoustic shadowing, 3) patella erosions: a cortical discontinuity or fragmentation of the inferior pole of the patella in the longitudinal or transverse plane (Figure 1A-C).

6

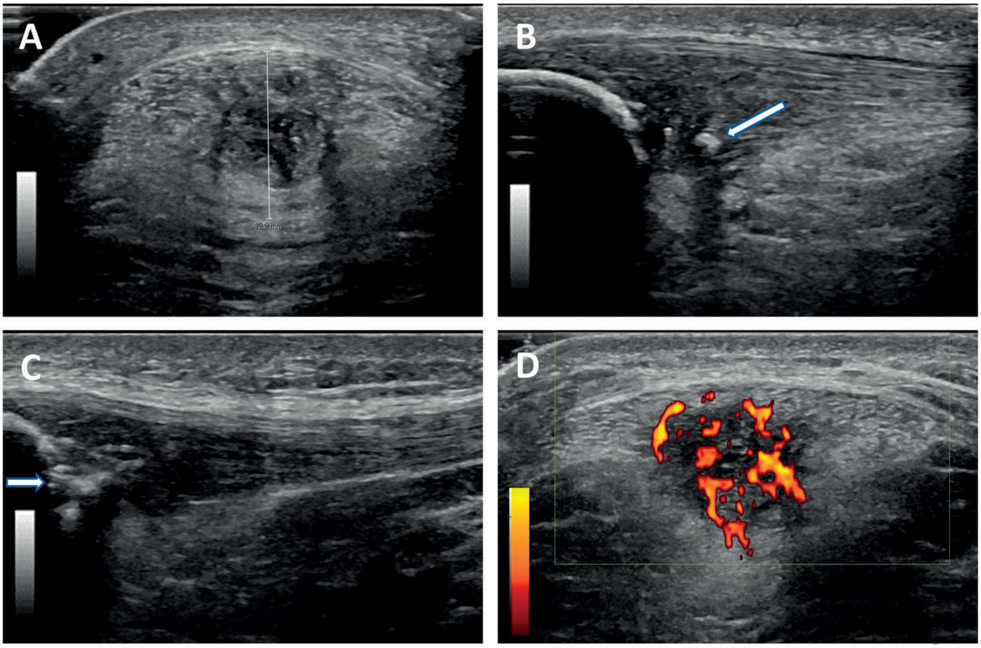


Figure 1. Grey-scale ultrasound (GSUS):

Anteroposterior anteroposterior (AP) thickness measured in the transverse plane (A), intratendinous calcifications identified in the sagittal plane (B), and patella erosions identified in the sagittal plane (C). Power Doppler ultrasound (PDUS): high level of Doppler flow (Modified Ohberg score, 4) assessed in the transverse plane (D). These images were from different athletes with PT.

PDUS was acquired in a transverse plane using a frequency of 10 MHz, a gain of 18 dB, a pulse repetition frequency of 1.0 kHz, and wall filters of 128 Hz. The box size of the region of interest (ROI) was 3.4 cm wide and 1.8 cm deep (Figure 1D). Participants were asked to extend both knees to increase the sensitivity of detecting (peri) tendinous blood flow[41]. The modified Ohberg score[46] was used to rate the tendinous Doppler signal using 5 scores (Table S1, Appendix A).

MRI

MRI was performed at baseline using a 3.0-T scanner (Discovery MR750, GE healthcare, Waukesha, WI, USA). All MRI scans were performed according to a previously described MRI protocol[8]. The protocol included the following pulse sequences: 3D fat-suppressed proton density-weighted sequence and axial fat-suppressed T2-weighted sequence (parameters for these 2 sequences were described in Table S2, appendix A).

The patellar tendon and surrounding tissues were evaluated for the following features: 1) focal tendon disruption: loss of fiber continuity on axial or sagittal images. In addition, greater than 50% focal tendon disruption was also collected in the axial view (one slice below the inferior pole of the patella) according to the previous literature[23], 2) IFP edema: any area of increased signal intensity within Hoffa fat pad, using a 4-point graded scale[16] (Table S1, Appendix A), 3) bone marrow edema: increased signal intensity within the bone marrow of the inferior pole of the patella[35], 4) deep infrapatellar bursitis: increased signal within the distended bursa[2] (Figure 2).

Imaging analyses

All US and MRI images were reviewed on our institutional picture archiving and communication system (PACS). An electronic caliper tool in PACS was used to determine AP patellar tendon thickness. All US images were analyzed by one musculoskeletal radiologist (EHGO, with 5 years of experience) and one sports physician (RJDV, with 3 years of experience). One musculoskeletal radiologist (YJF, with 11 years of experience) and the same sports physician (JD) analyzed all MRI images. Disagreements were mutually debated and solved with a third musculoskeletal radiologist (EHGO, with 20 years of experience). All these evaluators were blinded to the clinical outcomes at the time of image analysis.

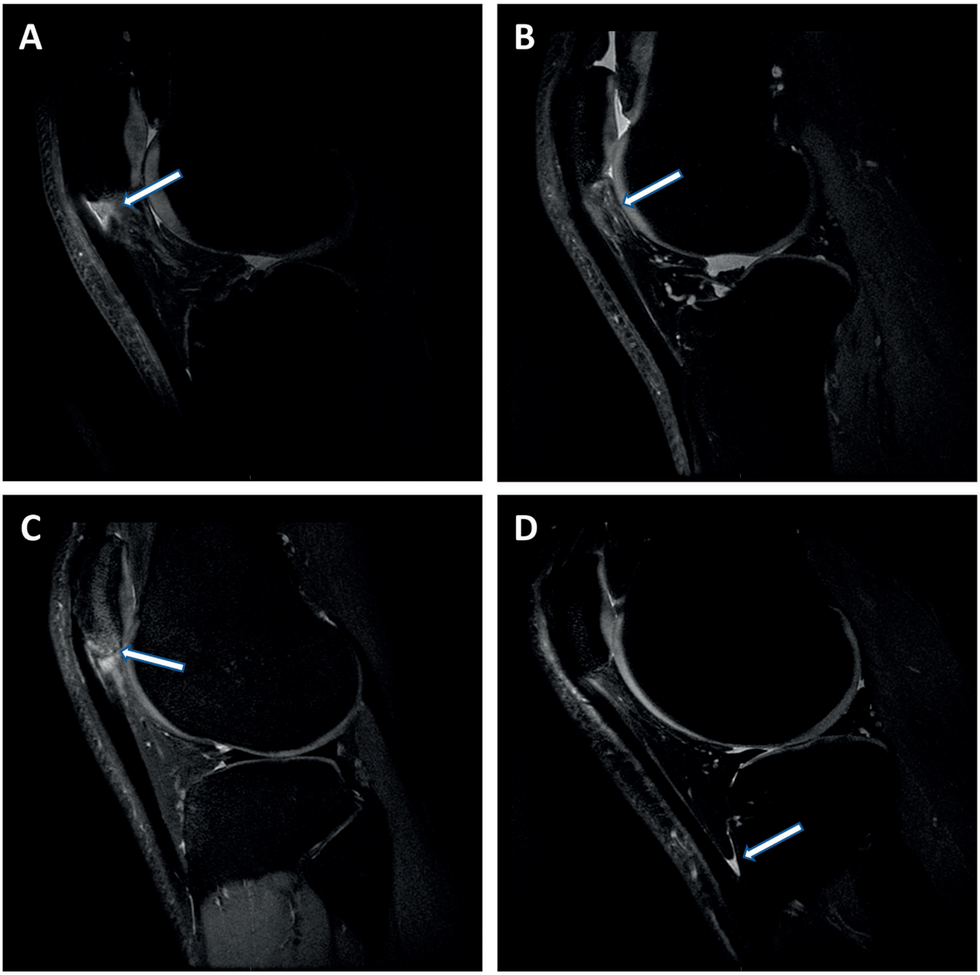


Figure 2. Sagittal, fat-suppressed proton density MR images:

Focal tendon disruption (A), grade 2 infrapatellar fat pad (IFP, or Hoffa fat pad) edema (B), the presence of bone marrow edema (C), and the presence of deep infrapatellar bursitis (D). These images were from different athletes with PT.

Statistical Analyses

Descriptive statistics for continuous data were presented as means with standard deviations for normally distributed data evaluated by visual check using histograms and also assessed by the Shapiro-Wilk test. Otherwise, medians and interquartile ranges were used.

We used linear mixed-effect models to evaluate the trajectory in VAS-SLDS and VISA-P scores over 24 weeks. The basic model included time and time-dependent covariates, including treatment allocation, sex, age, body mass index (BMI), symptom duration, and sports activity[3] as fixed effects. The baseline value of outcome and each imaging feature were added to the basic model to assess the prognostic value of imaging factors. The difference in the 24-week change of outcomes between individuals with and without imaging abnormalities was assessed by introducing an interaction term (time × imaging factor) into the model. The significance of this interaction was then tested using a likelihood ratio test (LRT), comparing models with and without the interaction term. All models included random intercepts and random slopes of time (constant up to 24 weeks) at the individual level. For AP thickness as a continuous predictor, we modelled non-linearity by fractional polynomials[45]. To facilitate clinical interpretation and enhance statistical power given sparsity in some ordinal categories, we dichotomized 3 predictors: AP thickness (diameter > 6mm[21]) for tendon thickening, Doppler flow (no to low [score 0 to 2] vs modest to high [score 3 to 4])[20], and IFP edema (no to low [grade 0 to 1] vs modest to high [grade 2-3]). Regression analysis was not performed for >50% focal tendon disruption due to its low prevalence (n = 3, 4%). We had also planned to assess the association between the presence of a hypoechoic region on ultrasound and increased signal intensity on MRI at the patellar tendon (i.e., features consistent with tendinopathy). However, all included athletes demonstrated these findings at baseline as it was an inclusion criterion, resulting in no variability within our study population. As a result, these features could not be evaluated as predictors in the regression models. We further evaluated the association between the number of imaging abnormalities (none, 1–2 abnormalities, 3–4 abnormalities, and 5 or more abnormalities) and the change in outcomes.

Linear mixed-effect models can handle missing outcomes as we assume the missing mechanism is missing at random. However, missing values in MRI features (Table 1) were imputed using multiple imputation. For the imputation model, we included the abovementioned covariates and auxiliary variables (details in Appendix A) and generated 20 imputed datasets. The results from each imputed dataset were pooled by Rubin's rules[9]. We repeated the primary analyses using complete cases as sensitivity analyses. The inter-rater reliability (IRR) for imaging findings among two raters was evaluated using Cohen's kappa for binary variables, Gwet's AC2 for ordinal outcomes[5], and interclass correlation coefficient (ICC) for continuous variables[4]. These values range from 0-1, and the responding cut-off values are described in Table S3, Appendix A.

We did not apply multiple testing correction due to the fact that these were exploratory analyses. Linear mixed-effect models were diagnosed by residual plots. All statistical analyses were conducted in R Studio version 2023.12.0 with 'mice', 'nlme', 'mitml', 'irr' and 'irrCAC' packages. A *P*-value of < 0.05 suggested statistical significance.

Results

Baseline characteristics and findings of imaging

A total of 76 PT athletes were studied (Table 1). Figure 3 shows the study flowchart. The response rates at 12 and 24 weeks were 84% and 88%, respectively.

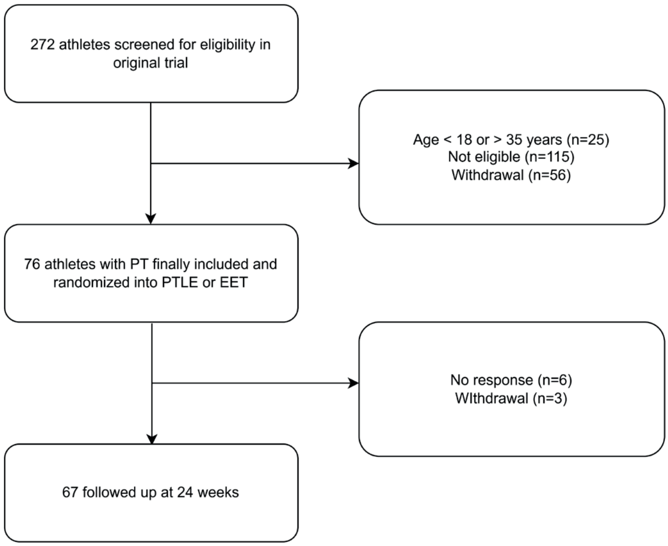


Figure 3: Flow chart of the study. Abbreviations:

PT, patellar tendinopathy; PTLE, progressive tendon loading exercises; EET, eccentric exercise therapy.

Among the 76 US examinations analyzed (100%), abnormal US features were identified in 26-78% of cases (Table 1). For MRI assessment, 72 out of 76 scans were available (95%, 4 missing data due to scanner unavailability caused by urgent clinical cases and time constraints). Among the 72 MRI scans analyzed, abnormal MRI features were demonstrated in 43-81% of athletes. All imaging features demonstrated good inter-rater reliability (Table S1, Appendix B).

Table 1. Baseline Demographic, Clinical, and Imaging Features in the Study Cohort (n=76)

	Value
Demographics	
Age, y, mean ± SD	24 ± 4
Male, n(%)	58 (76)
BMI, kg/m^2, mean ± SD	24 ± 3
Clinical factors	
Symptom duration, wk, median [IQR]	104 [49-208]
Sports activity level by CSAS, n(%)	
Level 1: 4-7 d/wk	17 (22)
Level 2: 1-3 d/wk	59 (78)
VAS-SLDS, (0-10), mean ± SD	5 ± 2
VISA-P score, (0-100), mean ± SD	55 ± 13
US	
AP thickness, mm , mean ± SD	9 ± 2
Thickening > 6 mm, n(%)	59 (78)
The presence of intratendinous calcification, n(%)	20 (26)
The presence of patella erosions, n(%)	24 (32)
Power Doppler flow, n(%)	
Level: no to low	25 (33)
Score 0	6 (8)
Score 1	2 (3)
Score 2	17 (22)
Level: moderate to high	51 (67)
Score 3	7 (9)
Score 4	44 (58)
MRI[a], n(%)	
The presence of focal tendon disruption	58 (81)
>50% focal tendon disruption	3 (4)
IFP edema	
Level: no to low	23 (32)
Grade 0	8 (11)
Grade 1	15 (21)
Level: moderate to high	49 (68)
Grade 2	19 (26)
Grade 3	30 (42)
The presence of bone marrow edema	31 (43)
The presence of deep Infrapatellar bursitis	37 (51)

Abbreviations: VAS-SLDS, Visual Analogue Score after single-leg decline squat; VISA-P, Victorian Institute of Sports Assessment-Patella; US, ultrasound; IFP, infrapatellar fat pad; AP, anteroposterior; CSAS, Cincinnati Sports Activity Scale; wk, week, d/wk, days per week; y, years; SD, standard deviation.

[a] The total sample size for athletes with MRI is 72.

Association between any imaging feature and 24-week change in pain and disability

The estimated mean change from baseline to 24 weeks was a decrease of 3 points (95% confidence interval [CI], 2 to 3, $P < 0.001$) in VAS-SLDS and an increase of 23 points (19 to 27, $P < 0.001$) in VISA-P score.

None of the evaluated imaging features was significantly associated with VAS-SLDS at 12 or 24 weeks (ß values range: -0.1 to 0.2; all P values > 0.1, Table 2). Moreover, interaction terms were not statistically significant (all $P_{interaction} > 0.1$), suggesting the presence of these imaging abnormalities at baseline did not significantly influence the 24-week change in VAS-SLDS compared to those without such features (Figure 4, Table S2, Appendix B).

Similar to VAS-SLDS, these imaging features were not significantly associated with VISA-P scores at 12 or 24 weeks (ß values range: -1.3 to 1.2; all P values > 0.1, Table 2), nor did they modify the course of in VISA-P scores over 24 weeks (all $P_{\text{interaction}} > 0.1$; Figure 5, Table S2, Appendix B).

Sensitivity analyses yielded similar results, although the precision of the estimates in cases with complete MRI data was generally lower (Table S3, Appendix B).

Association between the number of abnormal imaging features and 24-week change in pain and disability

When combining features from US and MRI, no abnormalities were observed in 5 out of 72 participants (7%). Among those with abnormalities (n = 67), 6 (9%) had 1 or 2 features, 23 (34%) had 3 or 4 features, and 38 (57%) had ≥ 5 features. PT athletes with a coexistence of multiple imaging abnormalities tended to report generally higher pain levels and greater disability during the follow-up time point. However, these differences were not statistically significant (Figure 6). The 24-week change in VAS-SLDS or VISA-P scores did not differ significantly according to the number of imaging abnormalities (all $P_{\text{interaction}} > 0.5$, figure 6).

Table 2. Results of Linear Mixed-Effect Models with 24-Week Change in Clinical Outcomes according to Imaging Features at Baseline

Imaging features	VAS-SLDS			VISA-P score		
	ß (95% CI)[a]	*P* value	*P* value for interaction[b]	ß (95% CI)[a]	*P* value	*P* value for interaction[c]
AP thickness, continuous	NA	NA	0.721	NA	NA	0.212
AP thickness, categorical			0.909	-1.3 (-4.64 to 2.08)	0.461	0.404
Tendon thickening vs normal	-0.0 (-0.61 to 0.56)	0.933				
Intratendinous calcifications			0.846			0.132
Present vs absent	0.1 (-0.51 to 0.63)	0.776		-0.2 (-3.41 to 2.95)	0.888	
Patellar erosions			0.159			0.283
Present vs absent	-0.1 (-0.60 to 0.44)	0.681		1.1 (-1.91 to 4.21)	0.468	
Power Doppler flow			0.305			0.354
Moderate to high vs no to low	0.1 (-0.37 to 0.65)	0.597		1.2 (-1.81 to 4.21)	0.442	
Focal tendon disruption			0.110			0.986
Present vs absent	0.0 (-0.58 to 0.64)	0.913		-1.3 (-4.65 to 2.11)	0.463	
IFP edema			0.442			0.115
Moderate to high vs no to low	0.1 (-0.39 to 0.59)	0.692		-0.5 (-3.46 to 2.48)	0.748	
Bone marrow edema			0.369			0.941
Present vs absent	0.2 (-0.22 to 0.70)	0.302		-1.2 (-3.94 to 1.54)	0.391	
Deep infrapatellar bursitis			0.145			0.512
Present vs absent	0.2 (-0.31 to 0.61)	0.515		-0.0 (-2.82 to 2.80)	0.996	

[a] The coefficients are derived from a linear mixed-effect model without an interaction term (imaging factor x time), representing the association between imaging factor and outcomes during follow-up (12 or 24 weeks).

[b] Likelihood ratio test for VAS-SLDS, comparing models with and without the interaction term (imaging factor x time), indicating the difference in 24-week change of VAS-SLDS between individuals with and without imaging factors.

[c] Likelihood ratio test for VISA-P score, comparing models with and without the interaction term (imaging factor x time), indicating the difference in 24-week change of VISA-P score between individuals with and without imaging factors.

Abbreviations: VAS-SLDS, Visual Analog Score after single-leg decline squat; VISA-P, Victorian Institute of Sports Assessment-Patella; US, ultrasound; IFP, infrapatellar fat pad; AP, anteroposterior; NA, not applicable.

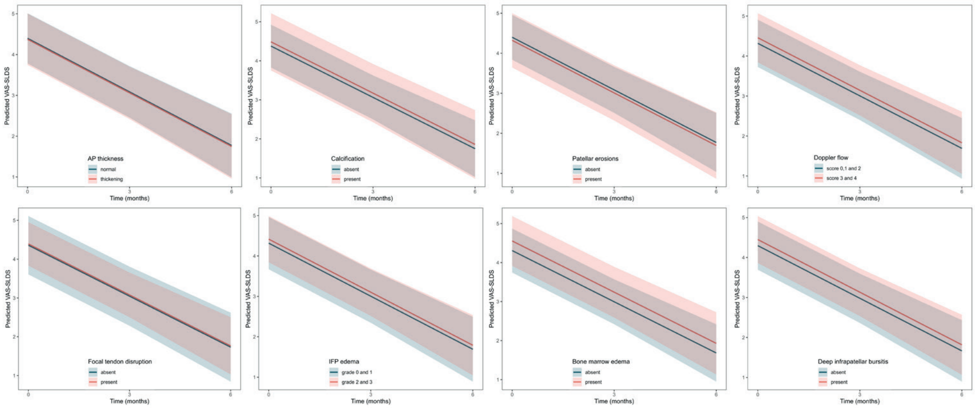


Figure 4. Effect plots visualize the 24-week decrease in VAS-SLDS according to the presence or absence of imaging features.

Parallel lines suggested that the extent of change in VAS-SLDS was similar between groups, indicating no prognostic value of any of these imaging features. The 95% CIs are calculated from linear mixed-effect models using shaded areas. Abbreviations: VAS-SLDS: Visual Analog Score after single-leg decline squat; IFP: infrapatellar fat pad.

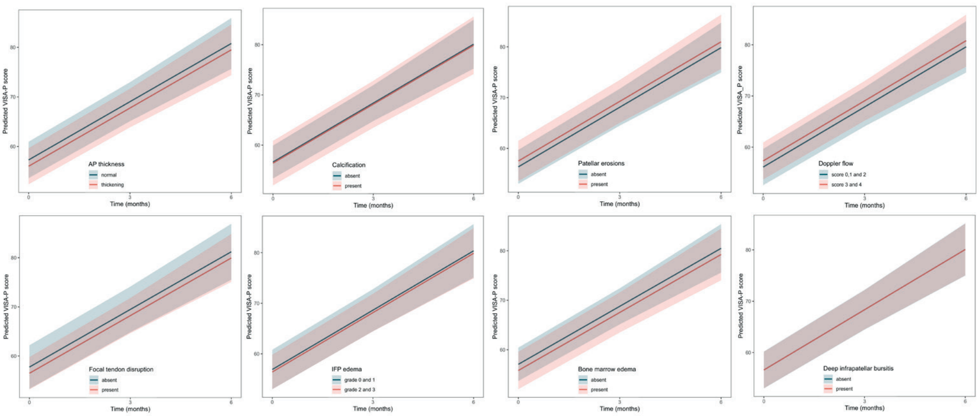


Figure 5. Effect plots visualize the 24-week increase in VISA-P score according to the presence or absence of imaging features.

Parallel lines suggested that the extent of change in VISA-P score was similar between groups, indicating no prognostic value of any of these imaging features. The 95% CIs are calculated from linear mixed-effect models using shaded areas. Abbreviations: VISA-P: Victorian Institute of Sports Assessment–Patella; IFP: infrapatellar fat pad.

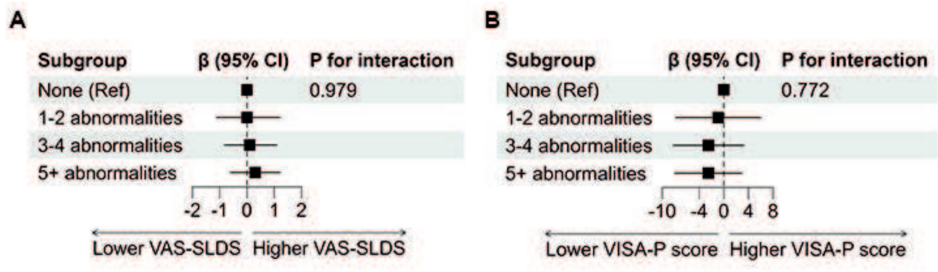


Figure 6. Forest plots visualize the association between the number of imaging abnormalities and outcomes during 24 weeks.

No significant interaction term indicates that the 24-week change in VAS-SLDS (A) or VISA-P score (B) did not differ according to the number of imaging abnormalities. Abbreviations: VAS-SLDS: Visual Analog Score after single-leg decline squat; VISA-P: Victorian Institute of Sports Assessment-Patella.

Discussion

To our knowledge, this is the first study investigating the prognostic value of conventional ultrasound (US) or MRI findings for predicting the 24-week change of pain and disability in PT athletes undergoing exercise treatment. Our findings revealed that, despite the high prevalence (26-81%) of abnormal imaging features in our study population, no single imaging feature was significantly associated with longitudinal change in clinical outcomes. Specifically, neither pain reduction measured by VAS-SLDS nor disability improvement assessed by VISA-P scores differed significantly between athletes with and without abnormal imaging features. Additionally, the number of abnormal imaging features did not significantly influence the 24-week changes in pain and disability. These results highlight that features identified by conventional imaging modalities may have limited utility in predicting recovery following exercise treatment for PT, despite the fact that many selected features are risk factors for developing PT[15,25,30]. These findings are crucial for optimizing patient education in the clinical setting.

The prognostic value of structural change in the patellar tendon

Previous research investigating the potential prognostic value of structural abnormalities in PT is limited to tendon thickening and partial tendon tears[23]. This is not surprising because an enlarged tendon is one of the most commonly observed structural changes in PT patients[39], with increased tendon thickness may correlate with the presence of tears[23]. To the best of our knowledge, only one previous study in PT has explored the relationship between these structural abnormalities and return to activity or sport after nonoperative treatment. Golman et al[23] reported that PT patients with tendon thickness (>11.5 mm) or substantial partial tear (>50% tear thickness) on MRI were a significant classifier for surgical treatment. However, this retrospective and lack of longitudinal design limit the ability to establish these imaging features as prognostic factors[24], as clinical decisions to pursue surgery may be influenced by clinician judgment, and the observed associations could be confounded by unmeasured variables. We did not find tendon thickness or the presence of a focal tear to be associated with a worse prognosis. Our findings are more reliably interpreted as prognostic, as we adhered to current guidelines for assessing prognostic factor[24,36,37] and accounted for potential non-linear relationships between tendon thickness and clinical outcomes in PT[40,45], which could induce erroneous inference[26]. While it is worthwhile to mention that compared to their study, our sample included few sufficient representations of severe cases (the mean tendon thickness was 8.8 mm (SD = 2.3), and only 3 cases (4%) had >50% tear thickness). All these 3 cases improved with exercise therapy and none needed surgery (Figure S1, Appendix B). Although conclusions cannot be drawn due to the very small sample size, these descriptive findings illustrate that partial tendon ruptures do not necessarily require surgical intervention.

The prognostic value of neovascularization

Although no prior research has directly evaluated the prognostic significance of neovascularization in PT patients performing exercise therapy, several studies have primarily focused on the effect of treatment targeting the eradication of neovascularization. A small-scale randomized trial[27] found borderline statistical significance for improved VISA-P scores at 4 months in PT patients receiving sclerosing injections aimed at reducing neovascularization compared to those receiving sham injections. Two prospective cohort studies reported general improvements in pain and disability scores following sclerotherapy[28] or electrocoagulation[11] of neovascularization. Taken together, these findings suggest that interventions reducing neovascularization could improve clinical outcomes, thus indirectly supporting a prognostic role. However, our results did not confirm the prognostic value of neovascularization assessed with PDUS. One possible explanation for this discrepancy could be that two out of these three previous studies[11,28] lacked control groups, making it difficult to determine whether the observed improvements were due to natural history or a true therapeutic effect from reducing neovascularization. Another reason might be our use of a dichotomized classification, categorizing neovascularization into “no to low” (scores 0 to 2) and “moderate to high” (scores 3 to 4). This simplified grouping may have limited our ability to discern the specific prognostic relevance of the presence versus absence of neovascularization. However, this methodological decision was aligned with our prior research[20] and was necessary to improve statistical model stability in our case, with only 8% of cases demonstrating no detectable Doppler signal and potential false positives[22] at low-grade neovascularization level (score 1, 3%). Future large-scale studies employing a more refined categorization of neovascularization might better clarify its specific prognostic impact on rehabilitation outcomes in PT.

The prognostic value of abnormalities in adjacent structures

There are no existing studies exploring the association between MRI-assessed abnormalities in adjacent structures and clinical outcomes in athletes with PT following exercise treatment. Only one prospective cohort (n= 30) found that preoperative IFP edema, or the coexistence of IFP and bone marrow edema in PT patients, was associated with worse pain and disability at a mean follow-up of 4 years after surgery[34]. Although it is debatable whether prognostic factors from the study population under different treatments can be directly compared, the previous study did not employ regression analyses, preventing its prognostic validity[37]. Considering they primarily included chronic or refractory cases who had failed at least 6 months of nonoperative treatment[34], it remains uncertain whether the reported associations reflect a true or independent prognostic relationship or are distorted by the underlying confounders such as chronicity of symptoms[43], greater functional impairments, and possible concurrent psychological or social factors[29] that may predict prognosis in tendinopathy.

Strengths and limitations

One of the key strengths of our study is that it was performed in accordance with the recommendations for assessing prognostic factors[37] under the PROGRESS framework[36]. By designing a prospective design with longitudinal outcomes, providing adjusted effects and using imputation methods to handle missing prognostic factors, we are able to reliably evaluate prognostic factors. Another strength is that we predefined two clinical outcomes, which are aligned within the core domains for tendinopathy[44]. The main limitation is that imaging features were assessed using binary classifications (presence or absence) rather than (semi-)quantitative measures, such the area of vascular signal[42]. Employing quantitative assessments might have provided a more in-depth understanding of the prognostic significance of these imaging features. Additionally, although the imaging features showed no prognostic value, the absence of a control group (e.g., patients receiving non-exercise treatment) limits our ability to assess whether these features modify the effect of exercise treatment (treatment effect modifier). Finally, due to a lack of established grading systems for classifying imaging abnormalities into specific subgroups, our categorization of coexisting imaging abnormalities into four levels (none, 1–2 abnormalities, 3–4 abnormalities, and ≥ 5 abnormalities) was data-driven. Therefore, the clinical relevance of these subgroupings should be interpreted with caution, and validation in future studies is warranted.

Clinical implications and future research

Despite the high prevalence of structural and vascular abnormalities on ultrasound and MRI in athletes with PT, these imaging features were not associated with the degree of clinical improvement following exercise therapy. Clinicians should be aware that these imaging findings do not reliably predict individual treatment outcomes. Patients should therefore be informed that the presence of these imaging abnormalities has a limited ability to determine their likelihood of recovery. This may help set realistic expectations and avoid unnecessary concern. Future research should include comparison groups receiving non-exercise-based interventions to assess whether imaging features act as treatment effect modifiers, and explore whether more refined or quantitative imaging metrics could better identify subgroups who respond differently to specific treatment approaches.

Conclusion

In this large prospective cohort study, we found no evidence that abnormalities assessed using conventional ultrasound or MRI were associated with 24-week changes in pain or disability among athletes with patellar tendinopathy undergoing exercise therapy. Although a substantial proportion of participants exhibited multiple coexisting imaging abnormalities, the number of abnormalities did not influence clinical outcomes over time. These findings indicate that these imaging features commonly used in clinical practice lack prognostic value in this context. Healthcare professionals should avoid relying on these imaging findings to predict prognosis. Patients can be informed that, despite the presence of structural or vascular abnormalities on imaging, they could still have a good chance of meaningful recovery with exercise-based treatment.

References

1. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, Magnusson SP. Clinical Outcomes and Tendon Structure at 3- to 4-Year Follow-up After Exercise-Based Treatment of Patellar Tendinopathy: A Prospective Study. Orthopaedic Journal of Sports Medicine. 2024;12(10):23259671241280192; PMID:39430115
2. Alqanatish JT, Petty RE, Houghton KM, et al. Infrapatellar bursitis in children with juvenile idiopathic arthritis: a case series. Clin Rheumatol. 2011;30(2):263-267; PMID:21088860
3. Barber-Westin S, Noyes FR. Assessment of Sports Participation Levels Following Knee Injuries. Sports Medicine. 1999/07/01 1999;28(1):1-10;
4. Benchoufi M, Matzner-Lober E, Molinari N, Jannot AS, Soyer P. Interobserver agreement issues in radiology. Diagn Interv Imaging. 2020;101(10):639-641; PMID:32958434
5. Benomar A, Zarour E, Létourneau-Guillon L, Raymond J. Measuring Interrater Reliability. Radiology. 2023;309(3):e230492; PMID:38085087
6. Breda SJ, de Vos R-J, Krestin GP, Oei EHG. Decreasing patellar tendon stiffness during exercise therapy for patellar tendinopathy is associated with better outcome. J Sci Med Sport. 2022; PMID:35094931
7. Breda SJ, Oei EHG, Zwerver J, et al. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. Br J Sports Med. 2021;55(9):501-509;
8. Breda SJ, Poot DHJ, Papp D, et al. Tissue-Specific T * Biomarkers in Patellar Tendinopathy by Subregional Quantification Using 3D Ultrashort Echo Time MRI. J Magn Reson Imaging. 2020;52(2):420-430; PMID:32108398
9. Buuren Sv. Flexible Imputation of Missing Data, Second Edition. 2nd Edition ed. 2018. https://doi.org/10.1201/9780429492259
10. Carr JC, Hanly S, Griffin J, Gibney R. Sonography of the patellar tendon and adjacent structures in pediatric and adult patients. AJR Am J Roentgenol. 2001;176(6):1535-1539; PMID:11373229
11. Carrozzo A, Shatrov J, El Helou A, et al. Ultrasound-guided electrocoagulation of neovascularisation for persistent patellar tendinopathy in athletes: a cohort study of 25 patients with a mean follow-up of 5 years from the SANTI Study Group. BMJ Open Sport Exerc Med. 2024;10(1):e001900; PMID:38500579
12. Challoumas D, Crosbie G, O'Neill S, Pedret C, Millar NL. Effectiveness of Exercise Treatments with or without Adjuncts for Common Lower Limb Tendinopathies: A Living Systematic Review and Network Meta-analysis. Sports Med Open. 2023;9(1):71; PMID:37553459
13. Challoumas D, Pedret C, Biddle M, et al. Management of patellar tendinopathy: a systematic review and network meta-analysis of randomised studies. BMJ Open Sport Exerc Med. 2021;7(4):e001110;
14. Culvenor AG, Cook JL, Warden SJ, Crossley KM. Infrapatellar fat pad size, but not patellar alignment, is associated with patellar tendinopathy. Scand J Med Sci Sports. 2011;21(6):e405-e411; PMID:21635562
15. Cushman DM, Stokes D, Vu L, et al. Ultrasound as a predictor of time-loss injury for the patellar tendon, Achilles tendon and plantar fascia in division I collegiate athletes. Br J Sports Med. 2025;59(4):241-248; PMID:39761995
16. De Smet AA, Davis KW, Dahab KS, Blankenbaker DG, del Rio AM, Bernhardt DT. Is there an association between superolateral Hoffa fat pad edema on MRI and clinical evidence of fat pad impingement? AJR Am J Roentgenol. 2012;199(5):1099-1104; PMID:23096185

17. de Vos R-J, van der Vlist AC, Zwerver J, et al. Dutch multidisciplinary guideline on Achilles tendinopathy. Br J Sports Med. 2021;55(20):1125-1134;
18. De Vries AJ, Koolhaas W, Zwerver J, et al. The impact of patellar tendinopathy on sports and work performance in active athletes. Res Sports Med. 2017;25(3):253-265; PMID:28391723
19. Deng J, Oosterhof JJ, Eygendaal D, Breda SJ, Oei EHG, de Vos R-J. Long-term Prognosis of Athletes With Patellar Tendinopathy Receiving Physical Therapy: Patient-Reported Outcomes at 5-Year Follow-up. Am J Sports Med. 0(0):03635465251336466; PMID:40356204
20. Deng J, Runhaar J, Breda SJ, Oei EHG, Eygendaal D, de Vos R-J. Do physical or imaging changes explain the effectiveness of progressive tendon loading exercises? A causal mediation analysis of athletes with patellar tendinopathy. J Sci Med Sport. 2024; PMID:39718487
21. Figueroa D, Figueroa F, Calvo R. Patellar Tendinopathy: Diagnosis and Treatment. The Journal of the American Academy of Orthopaedic Surgeons. 2016;24(12):e184-e192; PMID:27855131
22. Gatz M, Bode D, Betsch M, et al. Multimodal Ultrasound Versus MRI for the Diagnosis and Monitoring of Achilles Tendinopathy: A Prospective Longitudinal Study. Orthopaedic journal of sports medicine. 2021;9(4):23259671211006826; PMID:33912619
23. Golman M, Wright ML, Wong TT, et al. Rethinking Patellar Tendinopathy and Partial Patellar Tendon Tears: A Novel Classification System. Am J Sports Med. 2020;48(2):359-369; PMID:31913662
24. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med. 2006;144(6):427-437; PMID:16549855
25. Hjortshoej MH, Agergaard A, Larsen FK, et al. Determination of differences in ultrasound parameters for patellar tendons in males with unilateral patellar tendinopathy-An ancillary analysis of data from two randomized controlled trials. J Clin Ultrasound. 2024;52(5):548-557; PMID:38488675
26. Ho FK, Cole TJ. Non-linear predictor outcome associations. BMJ Med. 2023;2(1):e000396; PMID:36936259
27. Hoksrud A, Ohberg L, Alfredson H, Bahr R. Ultrasound-guided sclerosis of neovessels in painful chronic patellar tendinopathy: a randomized controlled trial. Am J Sports Med. 2006;34(11):1738-1746; PMID:16832128
28. Hoksrud A, Torgalsen T, Harstad H, et al. Ultrasound-guided sclerosis of neovessels in patellar tendinopathy: a prospective study of 101 patients. Am J Sports Med. 2012;40(3):542-547; PMID:22238058
29. Mc Auliffe S, Bisset L, Chester R, et al. ICON 2020-International Scientific Tendinopathy Symposium Consensus: A Scoping Review of Psychological and Psychosocial Constructs and Outcome Measures Reported in Tendinopathy Clinical Trials. J Orthop Sports Phys Ther. 2022;52(6):375-388; PMID:35647878
30. McAuliffe S, McCreesh K, Culloty F, Purtill H, O'Sullivan K. Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? A systematic review and meta-analysis. Br J Sports Med. 2016;50(24):1516-1523;
31. Millar NL, Silbernagel KG, Thorborg K, et al. Tendinopathy. Nat Rev Dis Primers. 2021/01/07 2021;7(1):1;
32. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375; PMID:19237405
33. Nutarelli S, da Lodi CMT, Cook JL, Deabate L, Filardo G. Epidemiology of Patellar Tendinopathy in Athletes and the General Population: A Systematic Review and Meta-analysis. Orthopaedic Journal of Sports Medicine. 2023;11(6):23259671231173659; PMID:37347023

34. Ogon P, Izadpanah K, Eberbach H, Lang G, Südkamp NP, Maier D. Prognostic value of MRI in arthroscopic treatment of chronic patellar tendinopathy: a prospective cohort study. BMC Musculoskelet Disord. 2017;18(1):146; PMID:28376759
35. Patel S. Primary bone marrow oedema syndromes. Rheumatology (Oxford). 2014;53(5):785-792; PMID:24080251
36. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS Med. 2013;10(2):e1001380; PMID:23393429
37. Riley RD, Moons KGM, Snell KIE, et al. A guide to systematic review and meta-analysis of prognostic factor studies. BMJ. 2019;364:k4597;
38. Rio EK, Mc Auliffe S, Kuipers I, et al. ICON PART-T 2019-International Scientific Tendinopathy Symposium Consensus: recommended standards for reporting participant characteristics in tendinopathy research (PART-T). Br J Sports Med. 2020;54(11):627-630; PMID:31519545
39. Rosen AB, Wellsandt E, Nicola M, Tao MA. Clinical Management of Patellar Tendinopathy. Journal of Athletic Training. 2022;57(7):621-631; PMID:34623447
40. Ryan M, Bisset L, Newsham-West R. Should We Care About Tendon Structure? The Disconnect Between Structure and Symptoms in Tendinopathy. J Orthop Sports Phys Ther. 2015;45(11):823-825; PMID:27136287
41. Smith E, Azzopardi C, Thaker S, Botchu R, Gupta H. Power Doppler in musculoskeletal ultrasound: uses, pitfalls and principles to overcome its shortcomings. J Ultrasound. 2021;24(2):151-156; PMID:32683646
42. van der Vlist AC, Veen JM, van Oosterom RF, van Veldhoven PLJ, Verhaar JAN, de Vos R-J. Ultrasound Doppler Flow in Patients With Chronic Midportion Achilles Tendinopathy: Is Surface Area Quantification a Reliable Method? J Ultrasound Med. 2020;39(4):731-739; PMID:31724758
43. van Rijn D, van den Akker-Scheek I, Steunebrink M, Diercks RL, Zwerver J, van der Worp H. Comparison of the Effect of 5 Different Treatment Options for Managing Patellar Tendinopathy: A Secondary Analysis. Clin J Sport Med. 2019;29(3):181-187; PMID:31033610
44. Vicenzino B, de Vos R-J, Alfredson H, et al. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. Br J Sports Med. 2020;54(8):444-451;
45. Visnes H, Bache-Mathiesen LK, Yamaguchi T, et al. Long-term Prognosis of Patellar Tendinopathy (Jumper's Knee) in Young, Elite Volleyball Players: Tendon Changes 11 Years After Baseline. Am J Sports Med. 2024:3635465241284648; PMID:39439271
46. Watson J, Barker-Davies RM, Bennett AN, et al. Sport and exercise medicine consultants are reliable in assessing tendon neovascularity using ultrasound Doppler. BMJ Open Sport Exerc Med. 2018;4(1):e000298; PMID:29527321

Supplementary material

Appendix A: methods

Multiple imputation

We performed multiple inputation using chained equations for missing MRI data. For the imputation model, we included confounders described in the method in the main text and auxiliary variables (sports participation in desired sport at baseline, sonographic imaging factors: tendont thickness, intratendinous calcification, patellar erosions and Doppler flow). The results from each imputed dataset were pooled based on Rubin's rule. For the likelihood ratio test, we used $D1$, considering we have a small sample size [1].

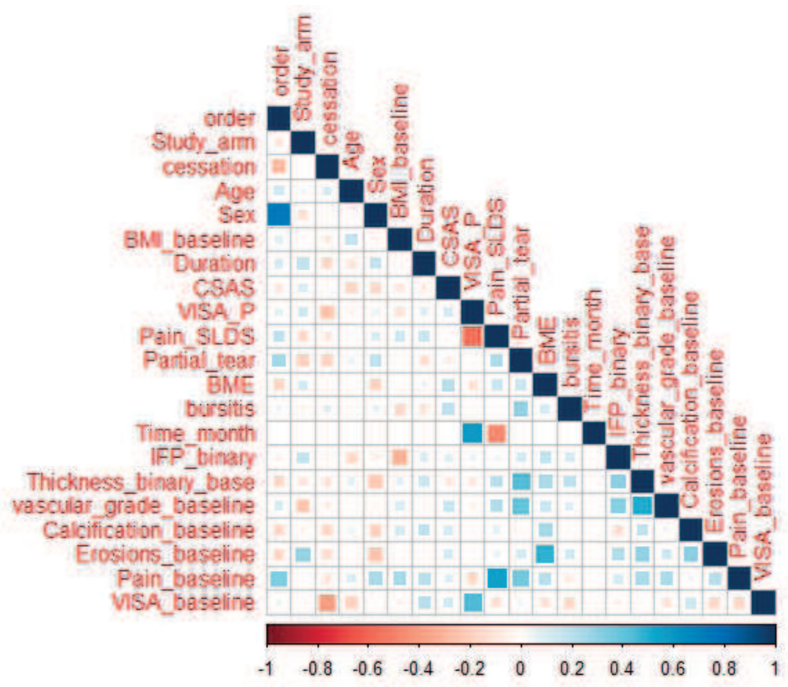


Figure S1: Correlation plot among variables in the imputation model

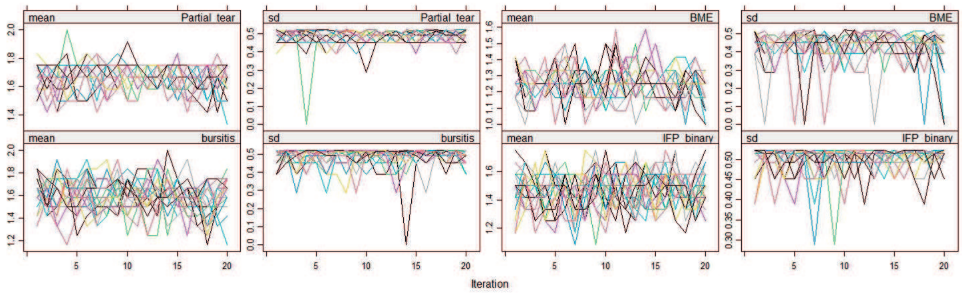


Figure S2. Convergence plot of 20 imputed dataset

[1] van Buuren, S. (2018). Flexible Imputation of Missing Data, Second Edition (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429492259

Table S1. Grey-scale ultrasound, power Doppler imaging, and MRI grading scheme

Imaging features	Measuring units or grading scores
Ultrasound	
Tendon AP thickness, continuous	
Tendon AP thickness, categorical	Normal (diameter ≤ 6 mm)
	Tendon thickening (diameter > 6 mm)
Intratendinous calcification(s)	Absent
	Present
Patella erosions	Absent
	Present
Doppler flow	No to low
	Score 0: no vessels
	Score 1+: one vessel posterior to the patella tendon
	Score 2+: one to two vessels throughout the tendon
	Moderate to high
	Score 3+: three vessels throughout the tendon
	Score 4+: four or more vessels throughout the tendon
MRI	
Focal tendon disruption	Absent
	Present
>50% focal tendon disruption	Absent
	Present
IFP edema	No to low
	Grade 0: no edema
	Grade 1: subtle increased signal intensity
	Moderate to high
	Grade 2: definite increased signal intensity
	Grade 3: marked increased signal intensity
Bone marrow edema	Absent
	Present
Deep Infrapatellar bursitis	Absent
	Present

Abbreviations: mm: millimeter; AP: anteroposterior; IFP: infrapatellar fat pad;

Table S2. Parameters of the magnetic resonance imaging sequences

Sequence	3D PD	Axial T2
Matrix	384 X 384	320 X 320
Scan plane	Sagittal	Axial
FOV (cm)	15.0	12.0
Resolution (mm)	0.4 X 0.4 X 1.0	0.375 X0.375
Slice Thickness (mm)	1.0	2.0
Number of Slices	120	32
TE (msec)	30.0	88.0
TR (msec)	1200.0	9378.0
Bandwidth (± kHz)	83.33	81.36
NEX	0.5	1
Fat saturation	Yes	yes

Abbreviations: PD: Proton density; FOV: field-of-view; TE: echo time; TR: repetition time; NEX: number of excitations

6

Table S3. Description of Inter-rater reliability calculations

Category	Statistics
Binary imaging variables	Cohen's k values were interpreted using the following cutoffs: <0 poor; 0.01 to 0.20 slight; 0.21 to 0.40 fair; 0.41 to 0.60 moderate; 0.61 to 0.80 substantial; and 0.81 to 1.00 almost perfect or 1.00 perfect [2].
Ordinal imaging variables	Gwet's AC2 values were interpreted using the following cutoffs: <0 poor; 0.01 to 0.20 slight; 0.21 to 0.40 fair; 0.41 to 0.60 moderate; 0.61 to 0.80 substantial; and 0.81 to 1.00 almost perfect or 1.00 perfect [2].
Continuous imaging variables	Interclass correlation coefficient (ICC) was interpreted using the following cutoffs: < 0.5 indicates poor reliability, ICC between 0.5 and 0.75 indicates moderate reliability, ICC between 0.75 and 0.9 indicates good reliability, and ICC > 0.90 indicates excellent reliability [3].

[2] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174; PMID:843571

[3] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15(2):155-163; PMID:27330520

Appendix B: results

Table S1. Inter-rater reliability (IRR) for imaging findings

Imaging features	IRR (95% confidence interval [CI])
Tendon AP thickness (mm)	0.83 (0.73-0.90)
The presence of intratendinous calcification	0.84 (0.69-0.96)
The presence of patella erosions	0.83 (0.69-0.95)
The level of Doppler flow	0.81 (0.75-0.88)
The presence of Focal tendon disruption	0.88 (0.72-1.00)
The presence of >50% focal tendon disruption	0.79 (0.77-1.00)
The presence of IFP edema	0.86 (0.82-0.91)
The presence of bone marrow edema	0.94 (0.85-1.00)
The presence of deep Infrapatellar bursitis	0.89 (0.78-0.97)

Abbreviation: IFP: infrapatellar fat pad

Table S2. Estimated mean difference from 24 weeks to baseline using linear mixed-effect models

Imaging features	Change in VAS-SLDS		Change in VISA-P score	
	Contrasts within subgroups Mean diff. (95% CI)	Contrasts between subgroups Mean diff. (95% CI)	Contrasts within subgroups Mean diff. (95% CI)	Contrast between subgroups Mean diff. (95% CI)
Tendon AP thickness				
Normal	-2.6 (-3.9 to -1.2)	0.1 (-1.4 to 1.6)	26.7 (17.7 to 35.7)	4.2(-6.1 to 14.5)
Tendon thickening	-2.7 (-3.4 to -1.9)		22.5 (17.5 to 27.5)	
Intratendinous calcifications		-0.1 (-1.6 to 1.4)		7.3 (-2.6 to 17.2)
Absent	-2.7 (-3.4 to -1.9)		25.3 (20.4 to 30.3)	
Present	-2.5 (-3.8 to -1.2)		18.0 (9.4 to 26.6)	
Patellar erosions		1.0 (-0.3 to 2.3)		-4.8 (-14.0 to 4.5)
Absent	-2.4 (-3.2 to -1.7)		21.9 (16.7 to 27.2)	
Present	-3.4 (-4.5 to -2.4)		26.7 (19.1 to 34.3)	
Power Doppler flow		0.8 (-0.5 to 2.0)		-4.1 (-13.2 to 5.0)
No to low (score 0-2)	-2.2 (-3.2 to -1.2)		20.8 (13.5 to 28.2)	
Moderate to high (score 3-4)	-3.0 (-3.8 to 2.3)		24.9 (19.6 to 30.3)	
Focal tendon disruption		1.2 (-0.3 to 2.7)		-0.1 (-10.7 to 10.5)
Absent	-1.7 (-3.0 to -0.4)		23.4 (14.1 to 32.7)	
Present	-2.9 (-3.6 to -2.2)		23.4 (18.5 to 28.4)	
IFP edema		0.5 (-0.8 to 1.9)		-7.1 (-16.2 to 2.0)
No to low (grade 0-1)	-2.3 (-3.4 to -1.2)		18.8 (11.4 to 26.2)	
Moderate to high (grade 2-3)	-2.8 (-3.6 to -2.0)		25.9 (20.6 to 31.2)	
Bone marrow edema		0.7 (-0.5 to 1.9)		0.3 (-8.5 to 9.1)
Absent	-2.4 (-3.2 to -1.6)		23.6 (18.0 to 29.2)	
Present	-3.0 (-4.0 to -2.0)		23.3 (16.5 to 30.0)	
Deep Infrapatellar bursitis		-0.9 (-2.2 to 0.3)		-2.8 (-11.6 to 5.9)
Absent	-3.1 (-4.1 to -2.2)		21.9 (15.5 to 28.3)	
Present	-2.2 (-3.1 to -1.3)		24.8 (18.9 to 30.7)	

Abbreviations: AP: anteroposterior; IFP: infrapatellar fat pad;

Table S3. Results of the linear mixed-effect model with 24-week change in clinical outcomes according to imaging factor at baseline using the complete dataset

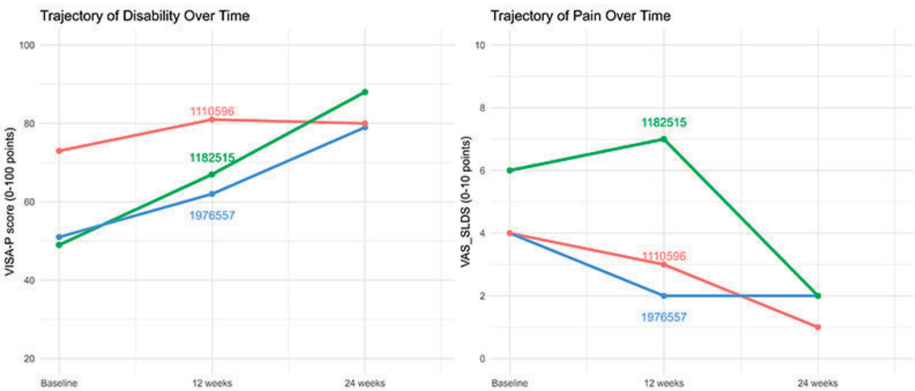
Imaging features	VAS-SLDS			VISA-P		
	ß (95% CI)	P value	P value for interaction[a]	ß (95% CI)	P value	P value for interaction[b]
Focal tendon disruption			0.096			0.915
Present vs absent	-0.0 (-0.67 to 0.61)	0.931		-0.9 (-4.45 to 2.64)	0.621	
IFP edema			0.455			0.076
Moderate to high vs no to low	-0.1 (-0.45 to 0.58)	0.803		-0.1 (-3.15 to 2.98)	0.956	
Bone marrow edema			0.526			0.736
Present vs absent	0.3 (-0.20 to 0.78)	0.253		-1.4 (-4.30 to 1.43)	0.333	
Deep infrapatellar bursitis			0.123			0.533
Present vs absent	0.2 (-0.31 to 0.63)	0.516		0.0 (-2.82 to 2.91)	0.976	

[a] Likelihood ration test for VAS-SLDS, comparing models with and without the interaction term (imaging predictor x time).

[b] Likelihood ration test for VAS-SLDS, comparing models with and without the interaction term (imaging predictor x time).

Abbreviations: VAS-SLDS: Visual Analogue Score after single-leg decline squat; VISA-P: Victorian Institute of Sports Assessment–Patella; IFP: infrapatellar fat pad.

Figure S1



Trajectories of VISA-P and VAS-SLDS scores from baseline to 24 weeks in three athletes with >50% focal tendon disruption at baseline, following exercise therapy. Each line represents an individual athlete's score trajectory; numbers indicate their study ID. None of the athletes underwent surgery during the 24-week follow-up period. Abbreviations: VAS-SLDS = Visual Analogue Scale after Single-Leg Decline Squat; VISA-P = Victorian Institute of Sport Assessment–Patella.

7

Chapter 7

General Discussion

This thesis covers two key domains of the management of patients with patellar tendinopathy (PT): enhancing patient education and understanding exercise-based treatment response. The first part aimed to investigate the performance of ChatGPT-4 in answering patient-centered questions related to PT, and to provide a 5-year prognosis of PT patients following structured and standardized exercise protocols. The second part of this thesis evaluated how findings from physical tests and conventional imaging relate to clinical outcomes. It specifically explored their potential role as mediators or prognostic indicators of exercise-based treatment effects.

Patient education for patellar tendinopathy

The role of ChatGPT in patient education

Comprehensive patient education is critical to manage individuals with PT.[1] It is associated with the success of loading-based treatment, empowers individuals with self-management, and thus may improve long-term prognosis.[2,3] As outlined earlier in the introduction of this thesis, there are still challenges in providing effective educational information about PT through in-person communication or traditional online platforms. In recent years, the rapid growth of large language models (LLMs) has opened new possibilities for how healthcare providers and patients engage with patient education.[4-6] Thus, there is an urgent need to equip stakeholders with knowledge about how LLMs, such as ChatGPT, perform in answering PT-related questions to ensure the benefits and pitfalls we are facing.

In **Chapter 2**, we benchmarked the output of ChatGPT-4 in response to queries related to PT management in a composite evaluation approach combining expert review and validated metrics that explicitly focus on patent education. We found that only around one out of three responses were accurate, and most were readable only by individuals with at least a college-level education, despite most of the responses being comprehensive, understandable, and actionable. This implies that this AI-powered tool is not yet suitable to serve as a standalone source of health-related information or advice for individuals with PT. A key concern is that patients may be exposed to misinformation that is conveyed with confidence, while they often lack the ability to critically assess accuracy and tend to trust human-like responses, such as those generated by ChatGPT.[7] In general, inaccuracies in these responses result from two main factors. First and foremost, LLMs like ChatGPT generate text by predicting words based on statistical patterns rather than engaging in true or human-like reasoning.[8,9] Consequently, their produced text may appear plausible and fluent but can be inaccurate or even fabricated, a phenomenon referred to as AI hallucinations.[10] Second, LLMs are trained on static datasets, which may be outdated and biased.[11] For instance, we found that ChatGPT-4 still recommends anti-inflammatory medications and eccentric exercises as cornerstone treatments in PT, despite evolving evidence that these treatments are not effective[1,12,13]. On the other hand, we feel healthcare providers could leverage this AI-powered tool to support PT patient education, provided that its accuracy can be verified. For instance, healthcare professionals could use ChatGPT-4 to assist in developing high-quality educational materials, addressing common shortcomings reported in current musculoskeletal resources.[14-17]

Clinicians in this field should equip themselves with this knowledge and remain informed, as individuals with PT may increasingly seek information through ChatGPT before consultations. Clinicians may encounter scenarios where patients bring up additional questions or topics based on prior exposure to ChatGPT or other LLMs' information. Proactively addressing misinformation is essential, as inaccurate information about treatment or prognosis can negatively impact adherence and outcomes (e.g., a patient may stop exercise therapy because they read that platelet-rich plasma (PRP) has a superior treatment effect). Engaging with patients on these topics may also foster trust and support shared decision-making.[18] We believe healthcare providers have a key role in verifying LLMs' generated content and guiding patients toward reliable sources of information, as most patients are receptive to providers' guidance on which internet resources to use.[19]

Setting realistic recovery expectations

The most important questions patients in general have, are related to recovery expectation or overall prognosis.[20] We believe this also matters in individuals with PT, supported by a substantial proportion of patient-centered questions related to prognosis (**Chapter 2**), and patients with PT who experienced contrasting expectations about rehabilitation outcomes.[18] When counseling individuals with patellar tendinopathy about their likelihood of symptom resolution or return to activity after rehabilitation, sports medicine clinicians face considerable uncertainty due to the wide variability in reported outcomes.[21-23]

To fill the gap, we revisited 76 patients with PT who received the current standard of care management consisting of patient education, activity modification, and structured exercise regimens (JUMPER study) and reported their overall prognosis at 5-year follow-up (**Chapter 3**).[24] We designed an online questionnaire containing several key domains of PT (global rating of change [GROC], symptoms, sports participation, disability, and quality of life), aiming to provide absolute risk and longitudinal change in these outcomes, which can be easily interpreted in decision-making.[25] We found that most participants (76%) could experience recovery and (71%) return to their desired sports at 5 years, while mild pain may remain. This indicates that a generally favorable long-term prognosis can be expected, which clinicians can use to support patient education, as communicating optimistic yet realistic expectations about recovery could foster belief in treatment and motivate participants to adhere to rehabilitation programs.[1,2,18]

It is important to acknowledge that approximately 30% of participants did not feel recovered at 5-year follow-up, and only 25% of participants reported being pain-free during sports participation. This raises the question of why certain subgroups experience a poorer prognosis, highlighting the need for personalized communication about recovery expectations and improved self-management strategies to reduce future healthcare utilization. To address this, we additionally investigated whether baseline patient-specific characteristics (sex, age, BMI, disability, and symptom duration) were associated with 5-year self-reported recovery, but none showed a significant association (**Chapter 3**). Our findings align with previous studies in Achilles tendinopathy, which also reported no associations between these factors and 5-year conservative treatment outcomes.[26,27] Notably, some of these variables were found to be associated with short-term outcomes. A pooled analysis of 3 randomized trials in individuals with PT following conservative treatment found that higher age and longer duration of symptoms were negatively associated with 3-month clinical improvement.[28] Similarly, in patients with rotator cuff tendinopathy, a systematic review identified that higher baseline disability and longer duration of symptoms were associated with unfavorable outcomes within 1 year.[29] These findings potentially explain the difficulty in identifying robust prognostic indicators for long-term recovery, as the effect of variables assessed at a single time point may be varied over time.[30] In addition, it is difficult to reliably assess variables like symptom duration due to potential recall bias. In patients with PT, the gradual or recurrent onset of symptoms makes it difficult to determine the true starting point, limiting the ability to assess accurate associations. Consistently, we also found no association between baseline physical tests (e.g., quadriceps muscle weakness) or imaging features (e.g., the presence of calcification or neovascularization on US) and 24-week change in disability in individuals with PT following exercise treatment (**Chapters 5 and 6**). This indicates that, while some of these factors have been identified as risk factors for PT, they were not associated with prognosis. Therefore, the common belief that we should address risk factors during rehabilitation is refuted by this study. Hence, patients should be informed that these proposed variables do not relate to long-term recovery. This also helps avoid unnecessary stress or anxiety.

Clinical and imaging indicators of exercise-based treatment response

Exploring mediators of PTLE versus EET

Progressive tendon loading exercises (PTLE) have been recommended as first-line treatment for individuals with PT.[1,31] This is supported by a high-quality randomized trial conducted previously in our group, where the 24-week difference between progressive tendon loading exercise (PTLE) and eccentric exercise therapy (EET) in disability (as measured by VISA-P) is evident[13]. Implementing these results in clinical practice will likely benefit from exploring underlying causal mechanisms (mediators) further, in other words, understanding how this intervention exerts its effect.[32] Mediation analysis is an ideal method to answer this question by decomposing the entire treatment effects into indirect effects carried through a given mediator and direct effects otherwise explained.[32,33] Such additional analyses can also extend resource investment in clinical trials, advance theory, optimize intervention, and contribute to the evidence of the management of PT.[32]

Since assuming a causal model and specifying the assumptions about this model are fundamental steps in mediation analysis, we spent a lot of time discussing the protocol.[32,33] We have to define what data and time points are needed from the trial to minimize bias in such analyses. We decided to explore whether the effect of PTLE compared to EET on pain-related disability and pain level during loading at 24 weeks mediated by improving physical (quadriceps muscle strength, range of ankle dorsiflexion and vertical jumping height) or imaging properties (tendon thickness or neovascularization on sonography) at 12 weeks, and if yes, by which magnitude? This pre-specified causal model was produced by scientific evidence and clinical-perspective knowledge, as well as assumptions that specifically refer to unmeasured confounders using causal directed acyclic graphs (DAGs).[34] Moreover, exploring these causal effects is highly clinically relevant, as it provides evidence on whether routinely performed physical tests and conventional US truly reflect treatment response in individuals with PT, given the frequent use of these assessments in both clinical practice and research.[13,35,36] Besides, we planned to perform sensitivity analyses evaluating the influence of unmeasured confounding and missing data to assess the robustness of the causal or statistical assumptions. We made these efforts to rigorously align with recommendations by the current guidelines of conducting a mediation analysis, making our results strongly reliable.[32,33]

In Chapter 4, we did not detect any significant indirect effect, suggesting the superior effect of PTLE over EET in pain-related disability and pain levels during loading was not mediated through changes in these selected factors. Sensitivity analyses yielded similar results. Previous studies have primarily investigated whether muscle strength mediates the effect of exercise treatment compared to non-exercise controls in individuals with gluteal tendinopathy, patellofemoral pain, and knee osteoarthritis.[37-39] Among these, only one study in individuals with knee osteoarthritis reported a statistically significant mediating effect of muscle strength.[39] However, the proportion of the treatment effect explained by muscle strength was minimal across all studies, at approximately 2%. In our study, we observed a slightly higher mediation effect of 5%, though it was not statistically significant, consistent with findings from the other two studies.[34,37,38] We consider differences in sample size to be a key factor contributing to these discrepancies. Three of the studies were secondary analyses of clinical trials, in which the sample size was calculated initially to detect treatment effects rather than mediation effects.[34,37,38] In contrast, one study used individual participant data pooled from multiple trials, resulting in a sample size exceeding 1,000 participants.[39] Although there are currently no established guidelines for sample size calculation specific to mediation analysis, we believe that larger sample sizes are necessary to detect statistically significant mediation effects with sufficient power, especially when the effect is minimal.[33] Besides, neurophysiological variables such as muscle activation or cortical inhibition might be more sensitive and are better at explaining changes in pain and function following PTLE.[40,41] This may be an area for future research.

No previous studies have investigated whether tendon thickness or neovascularization mediates the effects of exercise treatment in individuals with tendinopathy. Our study addressed this gap. To explain responses to therapeutic exercises, previous studies in tendinopathy have mainly examined whether reduction or normalization in altered tendon structure consistently parallel clinical improvement.[23,41-44] A systematic review in tendinopathy did not observe this trend in patients following eccentric exercise, though structural change may occur with heavy-slow resistance training.[44] Conversely, another systematic review in PT and Achilles tendinopathy reported a moderate correlation between reductions in tendon thickness and neovascularization and improved clinical outcomes, albeit based on small-sample studies.[36] A recent cohort study with long-term follow-up in PT observed reductions in tendon thickness and neovascularization following heavy or moderate-slow resistance exercises over 1 to 3.5 years, despite persistent symptoms.[23] These findings indicate that the time course for changes to tendon structure and clinical outcomes (pain and disability) following exercises may not parallel, Structural disorganization could persist after symptom resolution and some patients may show tendon restoration.[42]

To date, we conclude that physical tests and conventional ultrasound assessments have limited clinical utility either as potential treatment targets to improve outcomes or as a monitoring tool in individuals with PT undergoing PTLE. Further large-scale research comparing exercise vs non-exercise is needed to determine whether these assessments have any mediating effect.

Exploring prognostic factors for treatment response

Treatment responses in individuals with PT vary widely, with some cases achieving clinically meaningful change within weeks, whereas others may take months or continue to have symptoms for many years (**Chapter 3).**[13,35] Studying prognostic factors in this population helps clinicians predict who is at risk for a poor treatment response, supporting personalized decision-making process, and might even be useful as a modifiable target for further interventions, as well as facilitating personalized patient education, as stated above.[45]

We anticipated that there might be some prognostic value in physical (**Chapter 5**) or imaging (**Chapter 6**) factors given that they often appear as risk factors that are associated with the development of symptomatic PT.[46-48] For example, factors such as decreased ankle range of motion, increased patellar tendon thickness, or the presence of neovascularization on ultrasound have been associated with developing future pain at the patellar tendon.[46,48]

Our findings that conventional US or MRI features are not prognostic factors in PT are consistent with several studies of exercise programs for patients with Achilles tendinopathy.[26,49,50] We did also not identify tendon thickness or the presence of partial thickness tear (large areas of increased intratendinous signal intensity on MRI) as having a worse prognosis. However, in other scientific reports larger tendon thickness or >50% tear thickness were classifiers for surgical treatment in PT.[51] In patients with lateral elbow tendinopathy, the presence of partial-thickness or larger tears rather than tendon thickness was associated with poor response to nonoperative treatment in an univariate regression analysis.[52] A strength of our study was that we provided all patients with the same education about patellar tendinopathy, and we did not change the treatment allocation based on their imaging findings. Additionally, we performed adjusted analyses, allowing for a more reliable estimation of the prognostic value of imaging features.[53] A main disadvantage is that our data lacked sufficient representation of severe cases (e.g., >50% tear thickness) (**Chapter 6**), limiting our ability to detect a true association between the presence of a partial tear and patient-reported outcomes following exercise-based treatment.

The literature on the prognostic value of physical test findings of tendinopathies is limited. Only one prospective cohort study in AT identified that lower pain levels during loading tests (e.g., palpation and 10-hops test) were associated with greater improvement in 24-week disability.[54] However, the authors did not account for time-by-group interaction, limiting the interpretation of outcome changes over time and precluding direct comparison with our findings.

We believe that other strong prognostic factors, particularly psychological factors, may come into play to influence treatment response.[18,55] Constructs such as kinesiophobia, pain-related self-efficacy, pain beliefs, and fear-avoidance beliefs have recently been recommended as essential measures in tendinopathy research.[55] This is particularly relevant in patellar tendinopathy, where loading exercises often provoke pain, potentially leading patients to underload the tendon due to fear of further damage. Such insufficient tendon loading may prevent patients from achieving the clinical effectiveness of exercise-based treatment.[2] Further prognostic research is warranted to examine the role of these psychological constructs in exercise treatment response for patients with PT. It is important to note that our findings do not indicate a treatment effect modifier, which is defined as a baseline characteristic that is associated with a higher or lower treatment effect (intervention vs control group).[56] In our study, all patients received exercise-based treatment without a non-exercise group (control); therefore, the results can only be interpreted as showing no significant difference in 24-week outcome changes between patients with or without specific imaging features, indicating a prognostic, but not an effect-modifying, relationship.[45] We cannot determine whether patients with or without these features respond similarly to exercise compared to non-exercise treatments. Clarifying these definitions is essential, as they have different clinical implications.[57] Clinicians can utilize prognostic factors to stratify patients at risk of poor outcomes and enable more intensive follow-up or monitor for those individuals. Identifying treatment effect modifiers allows clinicians to select a specific treatment for subgroups of patients to benefit most, supporting a personalized treatment strategy. Although we agree that some of the prognostic factors can also be treatment effect modifiers, we were unable to verify this assumption in our case.[45]

Future perspectives

This thesis has extensively examined the management of patients with patellar tendinopathy (PT), with a particular focus on improving patient education and understanding the response to exercise-based treatment. We recommend that future research efforts prioritize the following areas.

Improve the performance of large language models (LLMs) in patient education

In the age of AI, the communication between patients and healthcare providers could be significantly transformed via AI-powered tools. LLMs such as ChatGPT show a great potential in assisting patient education. In this thesis, we only assessed the fundamental performance of ChatGPT in answering authorized questions. We encourage clinicians and patients to actively engage in prompt engineering to harness its full capacity to enhance the accuracy, clarity and readability of PT-related information.[58,59] An iterative process of refining prompts through joint feedback can be valuable—clinicians contribute insights on content accuracy and quality, while patients offer perspectives on satisfaction, relevance, and emotional resonance[59]. This approach based on LLMs like ChatGPT (a pre-train model) has a lot of advantages, including low costs, quick and easy for clinicians or non-technical users to optimize communication without needing to retrain or modify this model.

Instead, rather than relying on ChatGPT (a general LLM), sports medicine clinicians could even collaborate with engineers to develop PT-specific LLMs and conduct a randomized trial to investigate its effectiveness of improving patient education. We could train and evaluate a LLM using specific PT datasets regarding real-world communication between patients and clinicians via medical centers and online information. While such an approach is resource-intensive and requires substantial effort, it could enhance the model's relevance and effectiveness in supporting patient education and clinical decision-making within this domain.

We also recommend that future study designs should focus on determining the ability of LLMs to engage in interaction with patient-clinician conversations to mimic clinical practice. For instance, a study could be designed to investigate whether patients with PT using LLMs before the first consultation or after consultation under clinicians' supervision would improve adherence to exercise and positive beliefs, thus optimizing treatment outcomes. Nevertheless, it is always essential to carefully consider the ethical and safety issues associated with using AI in healthcare.[60]

Large-scale studies are needed for prognostic research

In this thesis, we determined the overall prognosis and prognostic factors using the data derived from a randomized trial. Despite the fact that we assessed the potential selection bias and adjusted treatment arm in the regression model, we acknowledge that eligibility criteria from randomized trials usually include filters (e.g., restricted in age, comorbidity) that exclude patients who are relevant to the prognostic questions of interest.[61] Moreover, the exercise treatment employed in randomized trials may be different from those utilized in the general practice.[20] Thus, further large-scale and prospective cohort studies need to validate our findings to increase the generalizability and statistical power.

We also suggest adding time-to-event data as one of the prognostic outcomes in PT-related research, as questions such as 'when to return to sport' or 'when to feel pain-free during sports' are of high importance to patients and clinicians. Another important consideration in prognostic studies is that they often involve long-term follow-up, in which missing data is common. Compared to binary outcomes assessed at fixed time points, time-to-event data captures the timing and trajectory of recovery, providing more informative and nuanced prognostic insights.

Causal inference to explain the treatment mechanism

Further randomized trials in PT should pre-specify the potential mechanisms of an intervention and plan to collect potential mediators at intermediate time points. This approach would support the advancement of theoretical understanding and enhance clinical implementation. Despite exercise being the first-line treatment for tendinopathy, the underlying mechanisms by which it leads to improvement remain poorly understood. We recommend the use of causal mediation analysis to investigate these mechanisms more rigorously. We hypothesize that several key domains may serve as potential mediators of exercise effectiveness in tendinopathy and warrant further exploration in future studies: 1) Tendon mechanical (e.g., change in tendon stiffness) or material properties (e.g, change in extracellular matrix): advanced imaging modalities such as shear wave elastography or ultrashort echo time (UTE)-MRI may help to quantify these parameters.[62,63] 2) Neurophysiological mechanisms: including factors such as muscle activation or cortical inhibition.[40,64,65] 3) Psychological mechanisms may contribute to treatment response through behavioral and perceptual pathways.[55]

Treatment effect modifier in randomized trials

There is an urgent need to shift from a one-size-fits-all treatment approach to identifying those specific patients who may benefit most or who hardly benefit from a certain intervention.[56] We would also advise clinicians to consider potential treatment effect modifiers during the design phase of randomized trials and explore these baseline factors in secondary analyses. However, individual trials are often underpowered to detect true effect modification (interaction term between modifier and treatment group).[66] Future meta-analyses or individual participant data (IPD) meta-analyses may be necessary to reliably identify and confirm treatment effect modifiers.

Conclusion

The purpose of this thesis was to enhance the management of patients with patellar tendinopathy (PT) by advancing patient education and improving the understanding of treatment response to exercise-based interventions.

Regarding patient education, we evaluated whether ChatGPT-4, as a large language model (LLM)-powered tool, could provide patient-centered questions related to PT. We found that it has the potential to generate comprehensive educational content, but it is not yet sufficiently accurate for patients to rely on independently when seeking PT-related information. These findings reinforce the importance of healthcare professionals remaining the primary source of patient education. Additionally, we found that patients with PT receiving care aligned with current standards can generally expect a favorable long-term prognosis. This evidence can assist clinicians in setting realistic recovery expectations and improving the quality of patient-centered communication during consultations.

Additionally, we investigated whether results from physical tests or conventional imaging findings could serve as mediators or prognostic indicators in patients with PT undergoing exercise-based treatment. Although none of the selected factors demonstrated a significant association with treatment mechanisms or prognosis, our findings contribute valuable insight into the clinical relevance and limitations of these commonly used assessments during the rehabilitation phase. This research highlights the need for continued evaluation of measurable clinical assessments to better inform prognosis and guide individualized treatment strategies in PT management.

Clinical implications of this thesis

Patient education

- Approximate three-quarters of responses from ChatGPT-4 related to PT were inaccurate, containing misinformation or outdated content delivered with high confidence. Clinicians should proactively inform patients of potential risks associated with the use of this tool, and may leverage this tool to support patient communication and reduce workload.
- In patients with PT receiving standardized exercise treatment, 76% recover and none worsen over 5 years; however, only 25% return to sport pain-free, suggesting that prolonged but acceptable symptoms may persist with sports participation.

Clinical and imaging indicators of exercise-based treatment response

- Improvements in quadriceps strength, ankle dorsiflexion, and jumping performance assessed by easy-to-perform physical tests, as well as tendon thickness or neovascularization on ultrasound, do not mediate better outcomes and should not be targeted to optimize the effect of exercise therapy in PT.
- Clinicians should not use initial findings from physical tests (that is: isometric quadriceps or hip abductors strength, lower extremity muscle flexibility, jump height, or pain during palpation, after jump tests, or single-leg decline squat) nor imaging (that is: sonographic tendon thickness, patellar erosions, calcification, neovascularization, or MRI-detected focal tendon disruption, bone marrow edema, infrapatellar edema or deep infrapatellar bursitis) to predict prognosis following exercise treatment in PT.

References

1. Rosen AB, Wellsandt E, Nicola M, Tao MA. Clinical Management of Patellar Tendinopathy. *Journal of Athletic Training*. 2022;57(7):621-631. http://doi.org/10.4085/1062-6050-0049.21.
2. Millar NL, Silbernagel KG, Thorborg K, Kirwan PD, Galatz LM, Abrams GD, Murrell GAC, McInnes IB, Rodeo SA. Tendinopathy. *Nat Rev Dis Primers*. 2021/01/07 2021;7(1):1. http://doi.org/10.1038/s41572-020-00234-1.
3. Hutting N, Johnston V, Staal JB, Heerkens YF. Promoting the Use of Self-management Strategies for People With Persistent Musculoskeletal Disorders: The Role of Physical Therapists. *J Orthop Sports Phys Ther*. 2019;49(4):212-215. http://doi.org/10.2519/jospt.2019.0605.
4. Campbell DJ, Estephan LE, Sina EM, Mastrolonardo EV, Alapati R, Amin DR, Cottrill EE. Evaluating ChatGPT Responses on Thyroid Nodules for Patient Education. *Thyroid*. 2024;34(3):371-377. http://doi.org/10.1089/thy.2023.0491.
5. Ayre J, Mac O, McCaffery K, McKay BR, Liu M, Shi Y, Rezwan A, Dunn AG. New Frontiers in Health Literacy: Using ChatGPT to Simplify Health Information for People in the Community. *J Gen Intern Med*. 2024;39(4):573-577. http://doi.org/10.1007/s11606-023-08469-w.
6. Davis R, Eppler M, Ayo-Ajibola O, Loh-Doyle JC, Nabhani J, Samplaski M, Gill I, Cacciamani GE. Evaluating the Effectiveness of Artificial Intelligence-powered Large Language Models Application in Disseminating Appropriate and Readable Health Information in Urology. *J Urol*. 2023;210(4):688-694. http://doi.org/10.1097/ju.0000000000003615.
7. Will ChatGPT transform healthcare? *Nat Med*. 2023;29(3):505-506. http://doi.org/10.1038/s41591-023-02289-5.
8. Hagendorff T, Fabi S, Kosinski M. Human-like intuitive behavior and reasoning biases emerged in large language models but disappeared in ChatGPT. *Nature Computational Science*. 2023/10/01 2023;3(10):833-838. http://doi.org/10.1038/s43588-023-00527-x.
9. Hua S, Jin S, Jiang S. The Limitations and Ethical Considerations of ChatGPT. *Data Intelligence*. 2024;6(1):201-239. http://doi.org/10.1162/dint_a_00243.
10. Farquhar S, Kossen J, Kuhn L, Gal Y. Detecting hallucinations in large language models using semantic entropy. *Nature*. 2024/06/01 2024;630(8017):625-630. http://doi.org/10.1038/s41586-024-07421-0.
11. OpenAI. Introducing ChatGPT. 2022. https://openai.com/blog/chatgpt. Accessed 03.22, 2024.
12. Deng J, Li L, Oosterhof JJ, Malliaras P, Silbernagel KG, Breda SJ, Eygendaal D, Oei EHG, de Vos R-J. ChatGPT is a comprehensive education tool for patients with patellar tendinopathy, but it currently lacks accuracy and readability. *Musculoskeletal Science and Practice*. 2025/04/01/ 2025;76:103275. http://doi.org/https://doi.org/10.1016/j.msksp.2025.103275.
13. Breda SJ, Oei EHG, Zwerver J, Visser E, Waarsing E, Krestin GP, de Vos R-J. Effectiveness of progressive tendon-loading exercise therapy in patients with patellar tendinopathy: a randomised clinical trial. *Br J Sports Med*. 2021;55(9):501-509. http://doi.org/10.1136/bjsports-2020-103403.
14. Win KT, Hassan NM, Oinas-Kukkonen H, Probst Y. Online Patient Education for Chronic Disease Management: Consumer Perspectives. *J Med Syst*. 2016;40(4):88. http://doi.org/10.1007/s10916-016-0438-0.
15. Schwarz I, Houck DA, Belk JW, Hop J, Bravman JT, McCarty E. The Quality and Content of Internet-Based Information on Orthopaedic Sports Medicine Requires Improvement: A Systematic Review. *Arthrosc Sports Med Rehabil*. 2021;3(5):e1547-e1555. http://doi.org/10.1016/j.asmr.2021.05.007.

16. Ghahremani JS, Ogi JE, Kody MT, Navarro RA. Readability of online patient education materials for shoulder instability surgery in English and Spanish. *J Shoulder Elbow Surg*. 2024;33(10):2220-2229. http://doi.org/10.1016/j.jse.2024.02.028.
17. Roberts H, Zhang D, Dyer GSM. The Readability of AAOS Patient Education Materials: Evaluating the Progress Since 2008. *The Journal of Bone and Joint Surgery American Volume*. 2016;98(17):e70. http://doi.org/10.2106/jbjs.15.00658.
18. Ruffino D, Sancho I, Alfonso M, Malliaras P. Beliefs and experiences of individuals with patellar tendinopathy following a rehabilitation program: A nested qualitative study. *Musculoskelet Sci Pract*. 2025;77:103311. http://doi.org/10.1016/j.msksp.2025.103311.
19. Koenig S, Nadarajah V, Smuda MP, Meredith S, Packer JD, Henn RF. Patients' Use and Perception of Internet-Based Orthopaedic Sports Medicine Resources. *Orthopaedic Journal of Sports Medicine*. 2018;6(9):2325967118796469. http://doi.org/10.1177/2325967118796469.
20. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KGM, Steyerberg EW, Roberts I, Schroter S, Altman DG, Riley RD. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ*. 2013;346:e5595. http://doi.org/10.1136/bmj.e5595.
21. Kettunen JA, Kvist M, Alanen E, Kujala UM. Long-term prognosis for jumper's knee in male athletes. A prospective follow-up study. *Am J Sports Med*. 2002;30(5):689-692.
22. Visnes H, Bache-Mathiesen LK, Yamaguchi T, Gilhuus HP, Algaard KRH, Hisdal E, Bahr R. Long-term Prognosis of Patellar Tendinopathy (Jumper's Knee) in Young, Elite Volleyball Players: Tendon Changes 11 Years After Baseline. *Am J Sports Med*. 2024:3635465241284648. http://doi.org/10.1177/03635465241284648.
23. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, Magnusson SP. Clinical Outcomes and Tendon Structure at 3- to 4-Year Follow-up After Exercise-Based Treatment of Patellar Tendinopathy: A Prospective Study. *Orthopaedic Journal of Sports Medicine*. 2024;12(10):23259671241280192. http://doi.org/10.1177/23259671241280192.
24. Ophey M, Koëter S, van Ooijen L, van Ark M, Boots F, Ilbrink S, Lankhorst NA, Piscaer T, Vestering M, den Ouden Vierwind M, van Linschoten R, van Berkel S. Dutch multidisciplinary guideline on anterior knee pain: Patellofemoral pain and patellar tendinopathy. *Knee Surg Sports Traumatol Arthrosc*. 2025;33(2):457-469. http://doi.org/10.1002/ksa.12367.
25. Vicenzino B, de Vos R-J, Alfredson H, Bahr R, Cook JL, Coombes BK, Fu SN, Gravare Silbernagel K, Grimaldi A, Lewis JS, Maffulli N, Magnusson SP, Malliaras P, Mc Auliffe S, Oei EHG, Purdam C, Rees JD, Rio EK, Scott A, Speed C, Akker-Scheek Ivd, Weir A, Wolf JM, Zwerver J. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. *Br J Sports Med*. 2020;54(8):444-451. http://doi.org/10.1136/bjsports-2019-100894.
26. van der Plas A, de Jonge S, de Vos RJ, van der Heide HJ, Verhaar JA, Weir A, Tol JL. A 5-year follow-up study of Alfredson's heel-drop exercise programme in chronic midportion Achilles tendinopathy. *Br J Sports Med*. Mar 2012;46(3):214-8.
27. Silbernagel KG, Brorsson A, Lundberg M. The majority of patients with Achilles tendinopathy recover fully when treated with exercise alone: a 5-year follow-up. *Am J Sports Med*. 2011;39(3):607-613. http://doi.org/10.1177/0363546510384789.
28. van Rijn D, van den Akker-Scheek I, Steunebrink M, Diercks RL, Zwerver J, van der Worp H. Comparison of the Effect of 5 Different Treatment Options for Managing Patellar Tendinopathy: A Secondary Analysis. *Clin J Sport Med*. 2019;29(3):181-187. http://doi.org/10.1097/jsm.0000000000000520.

29. Littlewood C, May S, Walters S. Epidemiology of Rotator Cuff Tendinopathy: A Systematic Review. *Shoulder & Elbow*. 2013;5(4):256-265. http://doi.org/10.1111/sae.12028.
30. Vittecoq O, Brevet P, Gerard B, Lequerre T. On difficulties to define prognostic factors for clinical practice in rheumatoid arthritis. *RMD Open*. 2024;10(3)http://doi.org/10.1136/rmdopen-2024-004472.
31. Marigi EM, Buckley P, Razi F, Abbas MJ, Jildeh TR, Camp CL, Krych AJ, Okoroha KR. Patellar Tendinopathy: Critical Analysis Review of Current Nonoperative Treatments. *JBJS Rev*. 2022;10(3) http://doi.org/e21.00168.
32. Cashin AG, McAuley JH, VanderWeele TJ, Lee H. Understanding how health interventions or exposures produce their effects using mediation analysis. *BMJ*. 2023;382:e071757. http://doi.org/10.1136/bmj-2022-071757.
33. Lee H, Cashin AG, Lamb SE, Hopewell S, Vansteelandt S, VanderWeele TJ, MacKinnon DP, Mansell G, Collins GS, Golub RM, McAuley JH, Localio AR, van Amelsvoort L, Guallar E, Rijnhart J, Goldsmith K, Fairchild AJ, Lewis CC, Kamper SJ, Williams CM, Henschke N. A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement. *JAMA*. 2021;326(11):1045-1056. http://doi.org/10.1001/jama.2021.14075.
34. Deng J, Runhaar J, Breda SJ, Oei EHG, Eygendaal D, de Vos R-J. Do physical or imaging changes explain the effectiveness of progressive tendon loading exercises? A causal mediation analysis of athletes with patellar tendinopathy. *J Sci Med Sport*. 2024;http://doi.org/10.1016/j.jsams.2024.12.006.
35. Agergaard A-S, Svensson RB, Malmgaard-Clausen NM, Couppé C, Hjortshoej MH, Doessing S, Kjaer M, Magnusson SP. Clinical Outcomes, Structure, and Function Improve With Both Heavy and Moderate Loads in the Treatment of Patellar Tendinopathy: A Randomized Clinical Trial. *Am J Sports Med*. 2021;49(4):982-993. http://doi.org/10.1177/0363546520988741.
36. Rabello LM, van den Akker-Scheek I, Brink MS, Maas M, Diercks RL, Zwerver J. Association Between Clinical and Imaging Outcomes After Therapeutic Loading Exercise in Patients Diagnosed With Achilles or Patellar Tendinopathy at Short- and Long-Term Follow-up: A Systematic Review. *Clin J Sport Med*. 2020;30(4):390-403. http://doi.org/10.1097/jsm.0000000000000624.
37. Mellor R, Kasza J, Grimaldi A, Hodges P, Bennell K, Vicenzino B. Mediators and Moderators of Education Plus Exercise on Perceived Improvement in Individuals With Gluteal Tendinopathy: An Exploratory Analysis of a 3-Arm Randomized Trial. *J Orthop Sports Phys Ther*. 2022;52(12):826-836. http://doi.org/10.2519/jospt.2022.11261.
38. Holden S, Matthews M, Rathleff MS, Kasza J, Vicenzino B. How Do Hip Exercises Improve Pain in Individuals With Patellofemoral Pain? Secondary Mediation Analysis of Strength and Psychological Factors as Mechanisms. *J Orthop Sports Phys Ther*. 2021;51(12):602-610. http://doi.org/10.2519/jospt.2021.10674.
39. Runhaar J, Holden MA, Hattle M, Quicke J, Healey EL, van der Windt D, Dziedzic KS, Middelkoop MV, Bierma-Zeinstra S, Foster NE, Group SOPA, Collaborative OATBE. Mechanisms of action of therapeutic exercise for knee and hip OA remain a black box phenomenon: an individual patient data mediation study with the OA Trial Bank. *RMD Open*. Aug 2023;9(3)
40. Rio E, Kidgell D, Purdam C, Gaida J, Moseley GL, Pearce AJ, Cook J. Isometric exercise induces analgesia and reduces inhibition in patellar tendinopathy. *Br J Sports Med*. Oct 2015;49(19):1277-83. http://doi.org/10.1136/bjsports-2014-094386.
41. Sprague AL, Couppé C, Pohlig RT, Cortes DC, Silbernagel KG. Relationships between tendon structure and clinical impairments in patients with patellar tendinopathy. *J Orthop Res*. 2022;http://doi.org/10.1002/jor.25262.

42. Ryan M, Bisset L, Newsham-West R. Should We Care About Tendon Structure? The Disconnect Between Structure and Symptoms in Tendinopathy. *J Orthop Sports Phys Ther*. 2015;45(11):823-825. http://doi.org/10.2519/jospt.2015.0112.
43. Färnqvist K, Pearson S, Malliaras P. Adaptation of Tendon Structure and Function in Tendinopathy With Exercise and Its Relationship to Clinical Outcome. *J Sport Rehabil*. 2020;29(1):107-115. http://doi.org/10.1123/jsr.2018-0353.
44. Drew BT, Smith TO, Littlewood C, Sturrock B. Do structural changes (eg, collagen/matrix) explain the response to therapeutic exercises in tendinopathy: a systematic review. *Br J Sports Med*. 2014;48(12):966-972. http://doi.org/10.1136/bjsports-2012-091285.
45. Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, Malats N, Briggs A, Schroter S, Altman DG, Hemingway H. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med*. 2013;10(2):e1001380. http://doi.org/10.1371/journal.pmed.1001380.
46. Sprague AL, Smith AH, Knox P, Pohlig RT, Grävare Silbernagel K. Modifiable risk factors for patellar tendinopathy in athletes: a systematic review and meta-analysis. *Br J Sports Med*. 2018;52(24):1575-1585. http://doi.org/10.1136/bjsports-2017-099000.
47. McAuliffe S, McCreesh K, Culloty F, Purtill H, O'Sullivan K. Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? A systematic review and meta-analysis. *Br J Sports Med*. 2016;50(24):1516-1523. http://doi.org/10.1136/bjsports-2016-096288.
48. Cushman DM, Stokes D, Vu L, Corcoran B, Fredericson M, Eby SF, Teramoto M. Ultrasound as a predictor of time-loss injury for the patellar tendon, Achilles tendon and plantar fascia in division I collegiate athletes. *Br J Sports Med*. 2025;59(4):241-248. http://doi.org/10.1136/bjsports-2024-109066.
49. Tsehaie J, Poot DHJ, Oei EHG, Verhaar JAN, de Vos RJ. Value of quantitative MRI parameters in predicting and evaluating clinical outcome in conservatively treated patients with chronic midportion Achilles tendinopathy: A prospective study. *J Sci Med Sport*. 2017;20(7):633-637. http://doi.org/10.1016/j.jsams.2017.01.234.
50. de Vos R-J, Weir A, Cobben LPJ, Tol JL. The value of power Doppler ultrasonography in Achilles tendinopathy: a prospective study. *Am J Sports Med*. 2007;35(10):1696-1701.
51. Golman M, Wright ML, Wong TT, Lynch TS, Ahmad CS, Thomopoulos S, Popkin CA. Rethinking Patellar Tendinopathy and Partial Patellar Tendon Tears: A Novel Classification System. *Am J Sports Med*. 2020;48(2):359-369. http://doi.org/10.1177/0363546519894333.
52. Clarke AW, Ahmad M, Curtis M, Connell DA. Lateral Elbow Tendinopathy:Correlation of Ultrasound Findings With Pain and Functional Disability. *Am J Sports Med*. 2010;38(6):1209-1214. http://doi.org/10.1177/0363546509359066.
53. Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, Hayden J, Collins GS, Debray TPA. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ*. 2019;364:k4597. http://doi.org/10.1136/bmj.k4597.
54. Mulder CFL, van der Vlist AC, van Middelkoop M, van Oosterom RF, van Veldhoven PLJ, Weir A, Verhaar JAN, de Vos R-J. Do physical tests have a prognostic value in chronic midportion Achilles tendinopathy? *J Sci Med Sport*. 2023;26(8):421-428. http://doi.org/10.1016/j.jsams.2023.06.014.
55. Stubbs C, McAuliffe S, Chimenti RL, Coombes BK, Haines T, Heales L, de Vos RJ, Lehman G, Mallows A, Michner LA, Millar NL, O'Neill S, O'Sullivan K, Plinsinga M, Rathleff M, Rio E, Ross M, Roy J-S, Silbernagel KG, Thomson A, Trevail T, van den Akker-Scheek I, Vicenzino B, Vlaeyen JWS, Pinto RZ, Malliaras P. Which Psychological and Psychosocial Constructs Are Important to Measure in Future Tendinopathy Clinical Trials? A Modified International Delphi Study With Expert Clinician/Researchers and People With Tendinopathy. *J Orthop Sports Phys Ther*. 2024;54(1):14-25. http://doi.org/10.2519/jospt.2023.11903.

56. Hingorani AD, Windt DAvd, Riley RD, Abrams K, Moons KGM, Steyerberg EW, Schroter S, Sauerbrei W, Altman DG, Hemingway H. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ*. 2013;346:e5793. http://doi.org/10.1136/bmj.e5793.
57. Sormani MP. Prognostic factors versus markers of response to treatment versus surrogate endpoints: Three different concepts. *Mult Scler*. 2017;23(3):378-381. http://doi.org/10.1177/1352458516676899.
58. Zaghir J, Naguib M, Bjelogrlic M, Névéol A, Tannier X, Lovis C. Prompt Engineering Paradigms for Medical Applications: Scoping Review. *J Med Internet Res*. 2024;26:e60501. http://doi.org/10.2196/60501.
59. Yan S, Knapp W, Leong A, Kadkhodazadeh S, Das S, Jones VG, Clark R, Grattendick D, Chen K, Hladik L, Fagan L, Chan A. Prompt engineering on leveraging large language models in generating response to InBasket messages. *J Am Med Inform Assoc*. 2024;31(10):2263-2270. http://doi.org/10.1093/jamia/ocae172.
60. Gerke S, Minssen T, Cohen G. Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A, Memarzadeh K, eds. *Artificial Intelligence in Healthcare*. Academic Press; 2020:295-336.
61. Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, McGinn T, Hayden J, Williams K, Shea B, Wolff R, Kujpers T, Perel P, Vandvik PO, Glasziou P, Schunemann H, Guyatt G. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ*. 2015;350:h870. http://doi.org/10.1136/bmj.h870.
62. Breda SJ, de Vos R-J, Krestin GP, Oei EHG. Decreasing patellar tendon stiffness during exercise therapy for patellar tendinopathy is associated with better outcome. *J Sci Med Sport*. 2022;http://doi.org/10.1016/j.jsams.2022.01.002.
63. Breda SJ, de Vos R-J, Poot DHJ, Krestin GP, Hernandez-Tamames JA, Oei EHG. Association Between T Relaxation Times Derived From Ultrashort Echo Time MRI and Symptoms During Exercise Therapy for Patellar Tendinopathy: A Large Prospective Study. *J Magn Reson Imaging*. 2021;54(5):1596-1605. http://doi.org/10.1002/jmri.27751.
64. Rio E, Kidgell D, Moseley GL, Gaida J, Docking S, Purdam C, Cook J. Tendon neuroplastic training: changing the way we think about tendon rehabilitation: a narrative review. *Br J Sports Med*. 2016;50(4):209-215. http://doi.org/10.1136/bjsports-2015-095215.
65. Vila-Dieguez O, Heindel MD, Awokuse D, Kulig K, Michener LA. Exercise for rotator cuff tendinopathy: Proposed mechanisms of recovery. *Shoulder & Elbow*. 2023;15(3):233-249. http://doi.org/10.1177/17585732231172166.
66. Tong G, Taljaard M, Li F. Sample size considerations for assessing treatment effect heterogeneity in randomized trials with heterogeneous intracluster correlations and variances. *Stat Med*. 2023;42(19):3392-3412. http://doi.org/10.1002/sim.9811.

Appendices

Summary

The overall aims of this thesis were to improve patient education in patellar tendinopathy by evaluating ChatGPT-4's performance in addressing patient-centered questions and providing updated long-term prognostic information, and to investigate whether physical and imaging parameters serve as mediators or prognostic factors in explaining treatment response following exercise-based therapy. **Chapter 1** presents the General Introduction of this thesis with an overview of the available literature on patellar tendinopathy with a focus on patient education and understanding exercise treatment responses.

Patient education

Large language models (LLMs) like ChatGPT have become a potential source for people seeking health-related information in an accessible and interactive manner. In **Chapter 2** we evaluated the performance of ChatGPT-4 in answering patient-care questions related to patellar tendinopathy (PT). Forty-eight questions were collected from online resources, patients with PT (n=10) and sports medicine experts (n=3). The assessment evaluated expert-rated accuracy and comprehensiveness, quality based on readability and patient education material criteria (understandability and actionability), and variability in responses over two months (including ChatGPT-4o) and across different PT-related terminologies. We found that 16 (33%) were rated accurate, while 36 (75%) were rated comprehensive. Most responses were written at a college reading level. The understandability and actionability were generally good. The meaning of responses and across terminologies from ChatGPT-4 were generally consistent. We concluded that this tool is currently not yet suitable to serve as a standalone patient education tool for PT patients.

In **Chapter 3** we conducted a prospective cohort study by re-visiting 76 patients with patellar tendinopathy (PT) from the JUMPER randomized trial, which compared progressive tendon-loading exercise (PTLE) and eccentric exercise therapy (EET). Five-year outcomes were assessed based on core tendinopathy domains, including self-reported recovery, pain, disability, quality of life, and sports participation. Among patients treated with structured exercise, 76% reported recovery and 71% returned to their desired sport, with sustained improvements in pain, disability, and quality of life. These findings indicate a generally favorable long-term prognosis, although persistent mild pain may remain. Baseline characteristics (sex, age, BMI, disability, and symptom duration) were not associated with long-term outcomes. Patients should be informed that these baseline variables do not predict long-term recovery.

Clinical and imaging indicators of exercise-based treatment response

The JUMPER study has shown that progressive tendon-loading exercise (PTLE) improved disability over 24 weeks more than eccentric exercise therapy (EET). To support clinical implementation, **Chapter 4** presents a mediation analysis exploring whether PTLE's benefits are mediated by changes in physical or imaging properties. We found no evidence that changes in quadriceps strength, ankle dorsiflexion, jumping performance, tendon thickness, or neovascularization mediated the superior effects of PTLE on disability or pain. These findings suggest that targeting these factors is unlikely to enhance PTLE outcomes, and alternative mechanisms should be considered in managing PT.

There are substantial variations in treatment response among patients with PT following therapeutic exercise. Identifying prognostic factors could help clinicians to stratify the risk of poor outcomes and guide decision-making. In **Chapter 5**, we collected baseline physical test results from patients of the JUMPER study and explored their prognostic value. We found muscle strength (quadriceps and hip abductors), muscle flexibility (quadriceps, hamstrings, soleus, and gastrocnemius), vertical jump height, and pain levels on palpation, after 3 jump trials, and after a single-leg squat were not associated with change in disability over 24 weeks.

Imaging is common assessed in clinical practice at baseline when managing PT. In **Chapter 6**, we also collected baseline ultrasound and MRI features from patients of the JUMPER study and explored their prognostic value. We found none of the selected imaging features (patellar tendon thickness, intratendinous calcifications, patellar erosions, and Doppler flow on US, and tendon fiber disruption, infrapatellar fat pad edema, bone marrow edema, and deep infrapatellar bursitis on MRI) were associated with change in disability over 24 weeks. We also demonstrated that there was no association between the total number of imaging features and change in disability over 24 weeks . Future prognostic research should focus on examining other factors to better understand the individuals' exercise treatment response in PT.

Chapter 7 presents the General Discussion with key findings of this thesis, followed by their interpretation in relation to the existing literature, and outlines future perspectives and clinical implications.

Samenvatting

De algemene doelen van dit proefschrift waren het verbeteren van de patiënt educatie in patella tendinopathie (PT) door de prestaties van ChatGPT-4 te evalueren bij het beantwoorden van patiëntgerichte vragen en het verschaffen van prognostische informatie op lange termijn, en te onderzoeken of klinische parameters en parameters op beeldvorming dienen als mediatoren of prognostische factoren bij het verklaren van het effect van oefentherapie. **Hoofdstuk 1** betreft de algemene introductie van dit proefschrift en geeft een overzicht van de beschikbare literatuur over PT met de nadruk op patiënt educatie en het begrijpen van de behandelrespons na oefentherapie.

Patiënt educatie

Large language models (LLM's) zoals ChatGPT zijn een potentiële bron geworden voor mensen die op zoek zijn naar gezondheidsgerelateerde informatie op een toegankelijke en interactieve manier. In **hoofdstuk 2** evalueerden we de prestaties van ChatGPT-4 bij het beantwoorden van vragen PT. Achtenveertig vragen werden verzameld van online bronnen, patiënten (n=10) en experts (n=3). Wij evalueerden de door experts beoordeelde nauwkeurigheid en volledigheid, de kwaliteit op basis van leesbaarheid en criteria voor patiënt voorlichtingsmateriaal (begrijpelijkheid en bruikbaarheid), en de variabiliteit in antwoorden over een periode van twee maanden (inclusief ChatGPT-4o) en voor verschillende PT-gerelateerde terminologieën. We ontdekten dat 16 antwoorden (33%) als accuraat werden beoordeeld, terwijl 36 antwoorden (75%) als uitgebreid werden beoordeeld. De meeste antwoorden waren geschreven op universitair leesniveau. De begrijpelijkheid en bruikbaarheid waren over het algemeen goed. De betekenis van de antwoorden en de terminologieën van ChatGPT-4 waren vrij consistent. Dit hulpmiddel is op dit moment niet geschikt om te dienen als zelfstandig hulpmiddel voor patiënt educatie voor PT.

In **hoofdstuk 3** voerden we een prospectief cohortonderzoek uit door 76 patiënten met PT uit het gerandomiseerde JUMPER-onderzoek, waarin progressieve peesbelastingsoefeningen (PTLE) en excentrische oefentherapie (EET) werden vergeleken, opnieuw te onderzoeken. Vijf-jaars uitkomsten werden beoordeeld op basis van kerndomeinen van tendinopathie, waaronder zelfgerapporteerd herstel, pijn, invaliditeit, kwaliteit van leven en sport participatie. Van de patiënten die werden behandeld met gestructureerde oefeningen, meldde 76% herstel en 71% keerde terug naar hun gewenste sport, met aanhoudende verbeteringen in pijn, invaliditeit en kwaliteit van leven. Deze bevindingen wijzen op een over het algemeen gunstige prognose op lange termijn, hoewel aanhoudende milde pijn kan blijven bestaan. Ongeveer een kwart rapporteerde geen herstel, en kenmerken die we hadden verzameld voor de start van de behandeling (geslacht, leeftijd, BMI, handicap en symptoomduur) waren niet geassocieerd met lange

termijn resultaten. Patiënten moeten worden geïnformeerd dat het niet goed mogelijk is om op basis van de genoemde karakteristieken het beloop van de klachten over lange termijn te voorspellen.

Klinische en beeldvorming parameters in relatie tot het effect van oefentherapie

Het JUMPER-onderzoek heeft aangetoond dat progressieve peesbelastingsoefeningen (PTLE) de functiebeperking gedurende 24 weken meer verbeterden dan excentrische oefentherapie (EET). Om de klinische implementatie te ondersteunen, werd in **hoofdstuk 4** een mediatie analyse verricht waarin werd onderzocht of de voordelen van PTLE worden gemedieerd door veranderingen in fysieke of beeldvorming parameters. De analyse vond geen bewijs dat verandering in quadriceps kracht, mate van dorsaalflexie van de enkel, sprong prestaties, peesdikte of neovascularisatie de superieure effecten van PTLE op invaliditeit of pijn medieerde. Deze bevindingen suggereren dat het onwaarschijnlijk is dat het veranderen van deze factoren de resultaten van PTLE verbetert en dat alternatieve mechanismen overwogen moeten worden als verklaring van het effect van PTLE.

Er zijn aanzienlijke verschillen in de respons op behandeling bij patiënten met PT na oefentherapie. Het identificeren van prognostische factoren zou clinici kunnen helpen bij het stratificeren van het risico op slechte uitkomsten en het sturen van de besluitvorming. In **hoofdstuk 5** verzamelden we fysieke testresultaten van patiënten van de JUMPER-studie en onderzochten we hun prognostische waarde. We ontdekten dat spierkracht (quadriceps en heupabductoren), spierflexibiliteit (quadriceps, hamstrings, soleus en gastrocnemius), verticale spronghoogte en pijnniveaus bij palpatie, na 3 sprongproeven en na een éénbenige squat niet geassocieerd waren met verandering in functiebeperking gedurende 24 weken.

We verzamelden in **hoofdstuk 6** baseline echografie- en MRI-kenmerken van patiënten van de JUMPER-studie en onderzochten hun prognostische waarde. We vonden dat geen van de geselecteerde beeldvorming parameters (patellapees dikte, intratendineuze verkalkingen, patella erosies en Doppler flow op echografie, en peesvezel disruptie, infrapatellair fat pad oedeem, beenmergoedeem en diepe infrapatellaire bursitis op MRI) geassocieerd was met verandering in functiebeperking na 24 weken. We toonden ook aan dat er geen verband was tussen het totale aantal beeldvormingskenmerken en verandering in functiebeperking na 24 weken. Toekomstig prognostisch onderzoek moet zich richten op het onderzoeken van andere factoren om de respons van patiënten met PT die oefentherapie doen beter te begrijpen.

Hoofdstuk 7 presenteert de algemene discussie met de belangrijkste bevindingen van dit proefschrift, gevolgd door een interpretatie in relatie tot de bestaande literatuur, en schetst toekomstperspectieven en klinische implicaties.

Phd portfolio

Courses	Year	ECTS
Master of Science in Health Sciences	2023-2025	70
Scientific integrity	2021	0.3
Biomedical Writing for PhD Candidates	2022	1.5
Student supervision		
Supervising Bachelor student Abdeslam Karbache	2025	3.0
Conferences and (oral) presentations		
Oral presentation Soup & Science meeting *Oral presentation at the Soup & Science meeting on the protocol of 5-year follow-up of JUMPER study*	2022	0.2
Oral presentation Soup & Science meeting *Oral presentation at the Soup & Science meeting on the mediation analysis in athletes with PT following exercise therapy*	2024	0.2
Oral presentation ADMIRE meeting *Oral presentation at ADMIRE meeting on prognostic value of conventional ultrasound in athletes with PT following exercise therapy*	2022	0.1
Oral presentation ADMIRE meeting *Oral presentation at ADMIRE meeting on the performance of ChatGPT in answering PT-related questions*	2024	0.1
Oral presentation Science day *Oral presentation at Science day on the prognostic value of physical test results in athletes with PT following exercise therapy*	2023	1
Oral presentation Science day *Oral presentation at Science day the on performance of ChatGPT in answering PT-related questions*	2025	1
Oral presentation at ISTS 2023 *The prognostic value of physical test results in athletes with PT following exercise therapy* Awarded best oral poster	2023	1
Poster presentation at ISTS 2023 *The prognostic value of conventional ultrasound findings in athletes with PT following exercise therapy*	2023	0.5

A

Courses	Year	ECTS
Poster Presentation at Symposium on Muscle-Tendon Function: Cell-Matrix-Cell Interplay *The prognostic value of physical test results in athletes with PT following exercise therapy*	2022	0.5
Poster Presentation European Congress of Radiology (ECR) *The prognostic value of conventional ultrasound findings in athletes with PT following exercise therapy*	2023	0.5
Oral Presentation ORS annual Meeting *3-minute PhD thesis pitch*	2025	1
Poster Presentation ORS annual Meeting *The mediation analysis in athletes with PT following exercise therapy*	2025	0.5
Poster Presentation ORS annual Meeting *Post-doc matching session*	2025	0.5
Other scientific activities		
Reviewer for Physical Therapy in Sport	2025	1.0
Reviewer for Research in Sports Medicine	2025	
Short-term Visit at University of Wisconsin Madison (Awarded by ORS Collaborative Exchange Grant)	2025	1.0 1.0
Total Workload in ECTS		84.9

List of publications

Association Between Physical Tests and Patients-Reported Outcomes in Athletes Performing Exercise Therapy for Patellar Tendinopathy: A Secondary Analysis of the JUMPER Study.
Deng J, Breda SJ, Eygendaal D, Oei EH, de Vos RJ. Am J Sports Med. 2023 Nov;51(13):3523-3532. doi: 10.1177/03635465231200241. PMID: 37815096.

ChatGPT is a comprehensive education tool for patients with patellar tendinopathy, but it currently lacks accuracy and readability.
Deng J, Li L, Oosterhof JJ, Malliaras P, Silbernagel KG, Breda SJ, Eygendaal D, Oei EH, de Vos RJ. Musculoskelet Sci Pract. 2025 Apr;76:103275. doi: 10.1016/j.msksp.2025.103275.. PMID: 39899928.

Do physical or imaging changes explain the effectiveness of progressive tendon loading exercises? A causal mediation analysis of athletes with patellar tendinopathy.
Deng J, Runhaar J, Breda SJ, Oei EHG, Eygendaal D, de Vos RJ. J Sci Med Sport. 2025 Jun;28(6):458-464. doi: 10.1016/j.jsams.2024.12.006. PMID: 39718487.

Long-term Prognosis of Athletes With Patellar Tendinopathy Receiving Physical Therapy: Patient-Reported Outcomes at 5-Year Follow-up.
Deng J, Oosterhof JJ, Eygendaal D, Breda SJ, Oei EHG, de Vos RJ. Am J Sports Med. 2025 Jun;53(7):1568-1576. doi: 10.1177/03635465251336466. PMID: 40356204.

Low socioeconomic status is associated with worse treatment outcomes in patients with Achilles tendinopathy.
Visser TS, Brul S, **Deng J**, Bonsel J, van Es E, Eygendaal D, de Vos RJ. Br J Sports Med. 2024 May 28;58(11):579-585. doi: 10.1136/bjsports-2023-107633. PMID: 38569849.

Letter to the Editor regarding Bullock et al: "Neck range of motion prognostic factors in association with shoulder and elbow injuries in professional baseball pitchers".
Deng J. J Shoulder Elbow Surg. 2025 Nov;34(11):e1131. doi: 10.1016/j.jse.2025.02.017. PMID: 40089012.

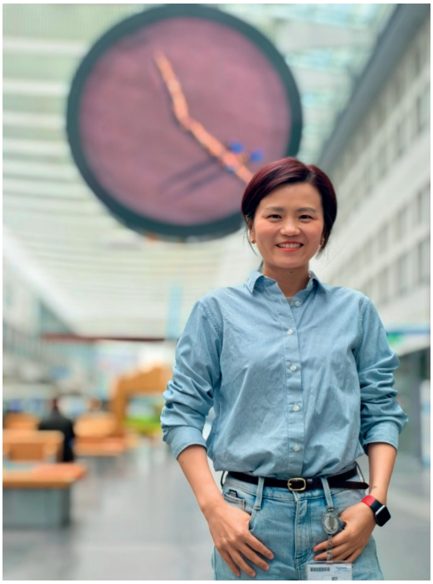
Conservative treatment of partial-thickness rotator cuff tears and tendinopathy with platelet-rich plasma: A systematic review and meta-analysis.
Xiang XN, **Deng J**, Liu Y, Yu X, Cheng B, He HC. Clin Rehabil. 2021 Dec;35(12):1661-1673. doi: 10.1177/02692155211011944. PMID: 33896214.

A

Exploring the digital landscape: A scoping review of Achilles tendinopathy education on public websites and in randomised controlled trials.
Malliaras P, Mallows A, McAullife S, Chimenti RL, Chen W, **Deng J,** Jiang J, Sharma S, Potter M, Smitheman H, Sancho I, Tavakkoli Oskouei S, Nicklen P, Bourke J, Fleagle T, Ruffino D, Silbernagel K, de Vos RJ. Clin Rehabil. 2025 Dec 8:2692155251397620. doi: 10.1177/02692155251397620. Epub ahead of print. PMID: 41358871.

Prognostic Value of Conventional Ultrasound and MRI Features for Clinical Outcomes in Athletes with Patellar Tendinopathy Following Exercise Therapy.
Deng J, Breda SJ, Fang YJ, Eygendaal D, de Vos RJ, Oei EH. Accepted, Sports Health: A Multidisciplinary Approach.

About the author



Jie Deng was born on September 28, 1991, in Dazhu, Sichuan Province, China. She earned her Bachelor's and Master's degrees in Clinical Medicine from Chongqing Medical University (2010–2018), followed by a residency in rehabilitation medicine at West China Hospital (2018–2021).

In 2021, Jie began her PhD at the Radiology & Nuclear Medicine and Departments of Orthopedics & Sports Medicine at Erasmus MC, Rotterdam, under the supervision of Prof. dr. Edwin Oei, Prof. dr. Denise Eygendaal, and Dr. Robert-Jan de Vos. Her research focuses on improving the management of athletes with patellar tendinopathy.

To support her research interests in biostatistics and epidemiology, she completed a second Master's degree in Health Sciences at the Netherlands Institute for Health Sciences (NIHES) in 2025.

After completing her PhD, Jie plans to continue academic research in sports medicine in the United States. In the long term, she is committed to a career as a sports physician, combining clinical care with research to address relevant clinical questions.

Acknowledgements

I feel truly fortunate to have experienced such a wonderful, fulfilling, and unforgettable PhD journey. These four years have witnessed my intellectual and personal growth. I have learned to think independently and critically, to embrace uncertainty, and to appreciate the subtle, harmonious balance between humans, technology, and nature. This journey has shaped my values, strengthened my perseverance, and deepened my understanding of what it means to contribute to knowledge with humility and responsibility.

None of these accomplishments would have been possible without the support, guidance, and encouragement of the many people who walked alongside me. Each challenge became manageable, each achievement more meaningful, thanks to their presence. For all of this, I am sincerely and profoundly grateful.

To my supervisors: Not everyone is lucky enough to be supported by three supervisors who are each distinguished experts in their respective fields. I am deeply grateful for this unique, multidisciplinary perspective, bringing together the expertise of a sports physician, a radiologist, and a sports surgeon, which has offered me an exceptional opportunity to learn from each of you. Your genuine, equal, and respectful collaboration has not only broadened my vision but has also shown me the true power of multidisciplinary teamwork.

To my daily supervisor, Robert-Jan: You were my very first mentor—patient, warm, and an inspiring role model. You guided me not only academically, shaping my scientific mindset, teaching me to think critically, and helping me understand what rigorous, meaningful research truly requires, but also personally, through your example of how a mentor supports, encourages, and empowers others. The mentorship I received from you has deeply influenced the kind of supervisor I hope to become in the future. I have been genuinely inspired by the way your work seamlessly bridges clinical practice and research, demonstrating how these two worlds can complement and strengthen one another. You taught me that the most important research questions are born from real clinical challenges, and that research should ultimately strive to improve patient care. You were also the one who first helped me articulate my long-term goal: to become a true clinician-researcher, someone who stands with one foot in clinical reality and the other in scientific discovery. Beyond academia, you have shown me what it means to be a well-rounded sports physician, with expertise across diverse disciplines, from cycling and skiing to skating and beyond. I have always admired the thoughtful balance you maintain between professional dedication and family life, setting an example of responsibility, kindness, and patience. Your mentorship has been unwavering, and at

many moments, you guided me with the care and wisdom of a father. Thank you for being a mentor in every sense, both within science and far beyond it.

To my supervisor, Edwin: You are an exceptional team leader, and I feel truly fortunate to be part of the ADMIRE group under your guidance. Our meetings are always thoughtfully structured and conducted with clarity and purpose, creating an environment where everyone feels engaged and supported. Your remarkable intelligence allows you to offer critical, insightful perspectives on all of our projects, even on topics far beyond your own specialty, which has continually broadened my scientific thinking. I deeply admire your mastery of time management: you switch effortlessly between tasks and responsibilities while consistently maintaining the same high standards of quality. Through your example, I have learned what effective and inspiring leadership truly looks like. You approach every situation with calmness, fairness, and humility, treating each member of the team with equal respect regardless of background or experience. This sense of equality and openness fosters a trusting and collaborative atmosphere that I have come to deeply appreciate. Most importantly, you are so approachable and easy-going. I have always felt comfortable sharing my thoughts, questions, and concerns with you without hesitation. Your patience, empathy, and down-to-earth personality have made a profound difference throughout my PhD journey.

To my exceptional supervisor, Denise: It is impossible to overstate how meaningful it has been to have you as one of my supervisors. As a female surgeon and the director of the department, you have inspired me time and time again through the way you lead, communicate, and work with both confidence and humility. Your professionalism, determination, and commitment to excellence have set a powerful example for everyone around you. I firmly believe that women benefit greatly from strong female role models, and your presence has provided exactly that, showing what is possible through resilience, clarity of vision, and genuine compassion. Seeing how you navigate demanding responsibilities while maintaining empathy, fairness, and grace has left a lasting impact on me. Your example is both profound and enduring, and will continue to shape the way I strive to grow in my own academic and clinical journey.

I would like to thank the **Graduate School** for providing me with the opportunity to participate in systematic and free courses in clinical research during my PhD trajectory. These courses have been both fundamental and practical, not only supporting me throughout every project during my PhD, but also providing the opportunity to become an independent researcher in the future. This program also marked my first English-based master's diploma and, more importantly, built my confidence in learning mathematics, biostatistics, and epidemiology, subjects that I have come to find extremely engaging and even exciting.

I am truly grateful to my Dutch friends and colleagues**: Huib, Martijn, Jessica**, **Daniek** and others, with whom I shared countless moments in our office, Ca-316. We had so many enjoyable lunch breaks, during which we exchanged ideas about cultural differences, investments, and personal life. Those conversations broadened my worldview and deepened my sense of openness and inclusiveness. Meanwhile, I'd like to thank my Chinese friends who work at Erasmus MC as well, **Qiaolin, Yifan, Jinchi, Delong** and **Wenjing**. You are the only ones who can truly feel me and root for each other from the heart, because we all know that this 4-year journal was not that easy and full of challenges. Finally, I would like to thank **Yijie Fang**, who joined our group during my final year of PhD studies. You are the most diligent, humble and sincere person I've ever met. It's so great to have you accompany me through this most stressful year and on many amazing trips.

感谢我的父母和小姨一家，每次回国都让我感到特别温暖和开心。

Lastly, but most importantly, I would like to thank my husband, **Lun Li**. You are the smartest and funniest person I have ever met. We have shared every moment together—whether meaningful or ordinary—and your presence fills my life with joy, peace, and strength. You are the one who truly understands me, who encourages me in moments of doubt, and who supports me unconditionally. Most exceptionally, you allow me the freedom to be entirely myself, without expectation or judgment. Without you, I would never have had the courage to go abroad, even though it has been my dream since childhood. Without you, I would not be who I am today. And without you, I would never have experienced how wonderful it feels to pursue something I genuinely love. Thank you for walking beside me with patience, humour, and unwavering faith—for everything seen and unseen. This achievement is as much yours as it is mine.