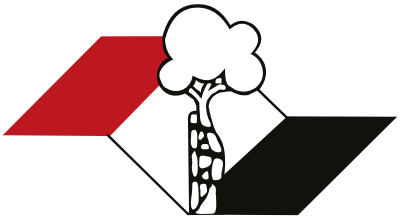
Indexed
PubMed and
PubMed Central
included
JCR (Journal Citation Reports®)



ISSN 1413-7852

Acta Ortopédica Brasileira

Volume 34 – Number 4 – Year 2026

ISSN 1413-7852

Acta Ortopédica Brasileira

Department of Orthopedics and Traumatology, Faculdade de Medicina da Universidade de São Paulo (DOT/FMUSP), São Paulo, SP, Brazil

Affiliated with Associação Brasileira de Editores Científicos

Indexed in PubMed, PubMed Central, Web of Science, JCR, Scopus Elsevier, SciELO, Redalyc (Red de Revistas Cientificas de America Latina y el Caribe,España y Portugal), LILACS (Latin America Health Science Literature) and DOAJ (Directory of open access journals).

EDITORIAL BOARD

Advisory Editor – Arthur Tadeu de Assis
Executive Editor – Ana Carolina de Assis

Administrative Editor – Atha Comunicação Editora
Logo creation – Caio Augusto de Souza Nery

Editorial coordination, creation, desktop publishing and graphic production Atha Comunicação e Editora - e-mail: 1atha@uol.com.br

SUMMARY

VOLUME 34 – Nº 4 – 2026

Original Article

DOI: http://dx.doi.org/10.1590/1413-785220263404e293218

ANATOMICAL STUDY OF THE INSERTION OF THE TRICEPS BRACHII TENDON IN THE OLECRANON THROUGH MAGNETIC RESONANCE IMAGING

ESTUDO ANATÔMICO DA INSERÇÃO DO TENDÃO DO TRÍCEPS BRAQUIAL NO OLÉCRANO POR RESSONÂNCIA MAGNÉTICA

José Renato Negrão[1], *Arnaldo Amado Ferreira Neto*[1], *Olavo Pires de Camargo*[1]

1. Universidade de Sao Paulo, Faculdade de Medicina, Hospital das Clinicas (HC-FMUSP), Departamento de Ortopedia e Traumatologia, Sao Paulo, SP, Brazil.

ABSTRACT

Introduction: The triceps brachii muscle (TBM) is the main muscle of the posterior aspect of the arm, occupying most of the extensor compartment. Proximally, it is composed of three heads: long, lateral, and medial, and presents a single tendon. Its insertion in the olecranon region remains controversial. Magnetic resonance imaging (MRI) assessment and detailed knowledge of this insertion may assist in orthopedic surgical reconstructions. Objective: To evaluate the anatomical aspect of the insertion of the triceps brachii tendon (TBT) into the olecranon regarding tendon insertion. Methods: *In vivo* MRI studies of patients were retrospectively evaluated. Two radiologists assessed 44 MRI scans from patients aged 20 to 50 years, with no previous triceps brachii tendon injuries or other TBT tendon pathologies. Results: There was agreement between the observers in the MRI scans evaluated, and the TBT insertion was found to be single. Conclusion: The triceps brachii tendon presented a single insertion into the olecranon. ***Level of Evidence III; Retrospective Comparative Study.***

Keywords: Magnetic Resonance Imaging; Tendons; Olecranon Process; Retrospective Studies.

RESUMO

Introdução: O músculo tríceps braquial (MTB) é o principal músculo da porção posterior do braço, preenchendo a maior parte do compartimento extensor. Na porção proximal, é composto por três cabeças: longa, curta e medial, apresentando um único tendão. Sua inserção na região do olécrano ainda é controversa. O estudo por ressonância magnética (RM) e o conhecimento pormenorizado dessa inserção auxiliam nas reconstruções cirúrgicas ortopédicas. Objetivo: Avaliar o aspecto anatômico da inserção do tendão do tríceps braquial (TTB) no olécrano em relação à inserção tendínea. Método: Estudos de RM de pacientes in vivo, avaliados de forma retrospectiva. Dois radiologistas avaliaram 44 exames de RM de pacientes com idade entre 20 e 50 anos, sem lesões prévias do tendão do tríceps braquial ou outras patologias tendíneas do TTB. Resultados: Houve uma concordância entre os observadores nas RM avaliadas e a inserção do TTB apresentou-se ser único. Conclusão: O tendão do tríceps braquial apresentou uma única inserção no olécrano. ***Nível de Evidência III; estudo Retrospectivo Comparativo.***

Descritores: *Imageamento por Ressonância Magnética; Tendões; Olécrano; Estudos Retrospectivos.*

Citation: Negrão JR, Ferreira Neto AA, Camargo OP. Anatomical study of the insertion of the triceps brachii tendon in the olecranon through magnetic resonance imaging. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 5. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

The triceps brachii muscle (TBM) is the main muscle in the posterior portion of the arm, filling most of the extensor compartment. The proximal portion is made up of three heads: long, short and medial, thus giving rise to its name.[1-2] Detailed knowledge of this insertion helps in orthopedic reconstruction, since the existing orthopedic literature is still very controversial.[1-3]

The triceps brachii muscle has three distinct muscle bellies and tendons. The current characteristics only present the description of a tendon inserting into the olecranon. The origin of the long head of the triceps is a flattened tendon of the infraglenoidal tubercle of the scapula, fusing above with the joint capsule of the shoulder. Its muscle fibers descend medially to the short head and superficially to the medial head, joining them to form a common tendon. [1-3]

The lateral head of the triceps arises from a flattened tendon arising from a narrow, linear, oblique ridge on the posterior surface of the shaft of the humerus and the lateral intermuscular septum, converges to the common tendon. After, the medial head is superimposed posteriorly by the short and long heads, exhibiting an extensive origin on the posterior surface of the humeral shaft below the radial

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Universidade de Sao Paulo, Faculdade de Medicina, Hospital das Clinicas (HC-FMUSP), Departamento de Ortopedia e Traumatologia, Laboratorio Professor Manlio Mario Marco Napoli, Sao Paulo, SP, Brazil.
Correspondence: José Renato Negrão. 333, Rua Dr. Ovidio Pires de Campos, Cerqueira Cesar, Sao Paulo, SP, Brazil. 05402-000. cjrnegrao@gmail.com

Article received on 01/07/2025 approved on 09/16/2025.
Handling Editor: Marcia Uchoa

groove, in the insertion plane of the teres major muscle, which is up to approximately 2.5 cm from the trochlea. Some muscle fibers reach the olecranon directly; others converge on the common tendon.[3]
The insertion tendon of the triceps brachii (TBT) begins in the medial portion of the muscle, having two blades: one superficial and the other deep. After the convergence of the muscle fibers, these two blades come together above the elbow and most fibers will insert into the superior surface of the olecranon.[3]
The study of TTB insertion in cadavers made a correlation between the anatomical dissection with the histological study, reporting that there were two insertions during dissection but only one during histological evaluation. [14-6]
Magnetic resonance imaging (MRI) is growing in clinical practice and being the key method for study the TBT in the olecranon because is a non-invasive modality and has high resolution between tissues. It is mainly used in cases of total or partial tendon ruptures. The objective of the study is to demonstrate in-vivo patients, using MRI, how the distal insertion of the TBT into the olecranon effectively occurs with the purpose of assist the orthopedist in tendon reconstruction.

MATERIALS AND METHODS

The study was carried out at the Institute of Orthopedics and Traumatology of the Hospital das Clínicas of the Faculty of Medicine of the University of São Paulo - FMUSP. In this study, 105 MRI exams of routine patients at the Institute of Orthopedics and the "Hospital das Clínicas of FMUSP were randomly analyzed. 61 exams were excluded from the study, due to the exclusion criteria used, such as age, that is, 44 exams. Factors such as age between 20 and 50 years and other clinical and surgical changes found in the studies meant that the number of exams was necessarily changed.
This elbow study is a retrospective analysis where they were carried out using 1.5 Tesla HDX® equipment (GE Healthcare, Waukesha, WI, USA) with a 1.5 Tesla HD Knee Transducer Coil (Channel TR Knee). The MRI protocol of the Radiology and Imaging Diagnostic Service of the Institute of Orthopedics and Traumatology adopted to study the insertion of the triceps brachii will be described below: a) Sagittal, axial and coronal planes; b) Fast spin echo technique; with T1 and T2 weighted images (matrix: 256, number of excitations (nex): 3, field of view (FOV) of 8 cm, thickness of 3.0 mm and gap of 0.5 mm). The images acquired during the exams were inserted into a data capture system (Pictures Archive Computer System Station – PACS). The positioning of patients, when performing the MRI examination, patients were positioned in the supine position with the arm above the head region.
All images were analyzed by two observers, radiologists with over 6 years of experience in imaging the musculoskeletal system. These evaluations were carried out independently and without prior knowledge of the results. The three sets of findings (two observers) will be used in data analysis and calculations of intra and inter observer correlations.
The images evaluated were used in the sagittal and axial planes in T1 and T2 weighted sequences. The images evaluated followed the plans produced by the Institute of Orthopedics and the "Hospital das Clínicas of FMUSP". These plans were used because they are the best sequences in relation to the information provided containing a more detailed analysis of the insertion of the triceps brachii tendon (TBT).

RESULTS

MRI elbow exams from 2012 to 2014 were eligible for the study. Therefore, from the 105 exams that were retrospectively included and after applying the study exclusion criteria, only a total of 44 exams were included, as shown in Figure 1.

According to Table 1, it shows that the exams evaluated came from MRI exams of patients aged between 20 and 50, half being male subjects. Still according to Table 1, it also shows that the age group between 41 and 50 had a higher frequency of imaging exams of the right elbow.
Each of these 44 exams was evaluated by two independent observers who classified the insertion of the triceps brachii tendon. Table 2 presents the results of these classifications where it shows that there was classification agreement between the two observers in 42 exams (95.4%). The Kappa agreement coefficient was calculated presenting a value equal to 0.847 (95% CI: 0.641; 1.000) and can be interpreted as a very good agreement.

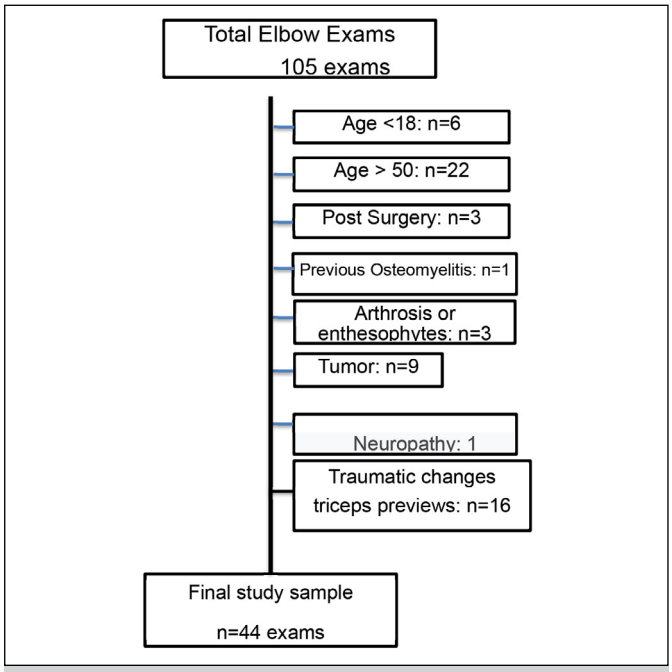


Figure 1. Study flowchart illustrating the exclusion criteria.

Table 1. Description of the demographic and clinical data of the 44 participating exams that made up the sample.

Age	**Number (%)**
Average ± standard deviation	37.9 ± 9.5
Minimum – maximum	20 – 50
Age Group	
20 to 30	11 (25.0)
31 to 40	13 (29.6)
41 to 50	20 (45.4)
Gender	
Male	22 (50.0)
Female	22 (50.0)
MRI	
Right Elbow	25 (56.8)
Left Elbow	19 (43.2)

Table 2. Distribution of the 44 classifications carried out by two independent observers regarding the insertion of the triceps brachii tendon.

Classification of tendon insertion		**Number (%)**
Observer 1	**Observer 2**	
Single insertion	Single insertion	35 (79.5)
Variation	Variation	7 (15.9)
Single insertion	Variation	1 (2.3)
Variation	Single insertion	1 (2.3)

Table 3 presents the results of the classification of tendon insertion by the two independent observers according to age group. It shows that, for elbows from MRI exams aged between 20 and 30, the two observers agreed that the single insertion was in all exams. For the age group between 31 and 40, there was agreement between the two observers in 12 exams (92.3%), while for the age group between 41 and 50, they agreed on 19 MRI exams (95%). The Kappa coefficient of agreement for the age group between 31 and 40, was equal to 0.755 (95%CI: 0.307 to 1.000) and for the age group between 41 and 50, the value obtained was equal to 0.875 (95%CI: 0.638; 1,000). These coefficients can be interpreted as showing good and very good agreement, respectively.
According to Table 4, it shows that the agreement occurred in 20 MRI female elbows exams (90.9%) and 22 MRI male elbows exams (100%). The Kappa agreement coefficient for females was equal to 0.771 (95% CI: 0.470 to 1.000), which can be interpreted as a good agreement. For males, it can be said that the agreement between the two observers was very good.
According to Table 5, it shows that the agreement occurred in 23 exams (92%) for the right-side elbows and in 19 exams (100%) for the left side. The Kappa agreement coefficient for the right side was equal to 0.702 (95% CI: 0.316 to 1.000), which can be interpreted as a good agreement. For the left side, the agreement between the two observers was very good.

Table 3. Distribution of the 44 classifications for triceps brachii tendon insertion carried out by two independent observers according to age group.

	Insertion Classification of the tendon insertion		Number(%)
Age Group	Observer 1	Observer 2	
20 to 30	Single Insertion	Single Insertion	11 (100.0)
31 a 40	Single Insertion	Single Insertion	10 (76.9)
	Variation	Variation	2 (15.4)
	Variation	Single Insertion	1 (7.7)
41 a 50	Single Insertion	Single Insertion	14 (70.0)
	Variation	Variation	5 (25.0)
	Single Insertion	Variation	1 (5.0)

Table 4. Distribution of the 44 classifications for triceps brachii tendon insertion carried out by two independent observers according to gender.

	Tendon Insertion Classification		Number (%)
Gender	Observer 1	Observer 2	
Female	Single Insertion	Single insertion	15 (68.2)
	Variation	Variation	5 (22.6)
	Single Insertion	Variation	1 (4.6)
	Variation	Single Insertion	1 (4.6)
Male	Single Insertion	Single Insertion	20 (90.9)
	Variation	Variation	2 (9.1)

Table 5. Distribution of the 44 classifications for insertion of the triceps brachii tendon carried out by two independent observers according to the side of the elbow.

	Tendon Insertion Classification		Number (%)
Elbow Side	Observer 1	Observer 2	
Right	Single Insertion	Single Insertion	20 (80.0)
	Variation	Variation	3 (12.0)
	Single Insertion	Variation	1 (4.0)
	Variation	Single Insertion	1 (4.0)
Left	Single Insertion	Single Insertion	15 (79.0)
	Variation	Variation	4 (21.0)

Statistical Analysis

A descriptive analysis was performed where categorical variables were summarized by the number (*n*) and percentage (%) and non-categorical variables as an average ± standard deviation, minimum and maximum values.
Agreement between the two observers was assessed by calculating the observed agreement (proportion of exams with the same classification by both observers) and the Kappa-Cohen agreement coefficient (κ) and its respective 95% confidence interval (95%CI). The Kappa coefficient varies between 0 and 1, where higher values indicate greater agreement between observers. The degree of Kappa agreement was interpreted as poor (Kappa ≤ 0.20), fair (Kappa between 0.21 and 0.40), moderate (Kappa between 0.41 and 0.60), good (Kappa between 0.61 and 0.80) and very good for Kappa values between 0.81 and 1.00 (Altman, 1991).
Statistical analysis was done using Stata/MP 18.0 (Stata-Corp, 2023. College Station, TX: Stata Corp LLC).
Altman DG (1991) Practical statistics for medical research. London: Chapman and Hall.

DISCUSSION

In the literature that was evaluated, several studies showed findings of the triceps brachii in cadavers together with dissections and histological studies. However, this is the first study that evaluates the anatomical findings of the triceps brachii retrospectively from in-vivo examinations, according to Figure 2.
The MRI study of the insertion of the triceps brachii tendon has been the key topic of several studies. The main objective of this study is to assist orthopedists and radiologists in the surgical correction of the triceps brachii tendon and to expand the analysis of its insertion. This finding should be considered in clinical practice to achieve greater assurance and more positive results in reconstruction surgery for triceps brachii tendon injuries.
This study was carried out using MRI as this method presents better resolution between tissues and does not use ionizing radiation. The radiography (x-ray) study does not present details of soft tissues that are necessary to enable the evaluation of the tendon. Computed tomography study would not be able to participate in this study because it also does not present fine necessary details and both methods use ionizing radiation. The ultrasound study presents tissue details without using ionizing radiation but, it is an

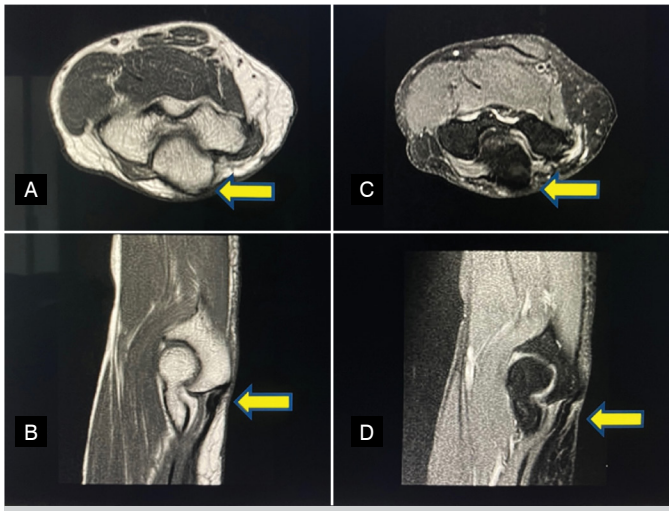


Figure 2. (A) and (B) T1 Sequence of axial and sagittal MRI scans show the distal triceps tendon insertion in the olecranon (yellow arrows); (C) and (D) T2 Sequence of axial and sagittal show the distal triceps insertion in the olecranon (yellow arrows).

examiner-dependent method and has limitations such as the low reproducibility of the methodology. [8-12]

Traditional anatomical descriptions indicate three heads of the triceps brachii muscle called the long head, the lateral head, and the medial head, which insert into the olecranon. The long head originates basically from the infraglenoidal tubercle of the scapula and inferior glenoumeral joint capsule. The lateral head has three points of origin: the posterior surface of the humerus between the insertion of the teres minor tendon and the superior surface of the spiral groove, the lateral edge of the humerus, and the lateral intermuscular septum. The head originates from the posterior surface of the humerus, distal to the spiral groove and medial aspect of the intermuscular septum. The medial and lateral heads serve only as extensors of the elbow, while the long head assists in adduction and extension of the glenohumeral joint.[12-15]

The triceps brachii tendon has a different characteristic from other tendons, as it originates from three muscle bellies. These muscle bellies originate from specific bone locations as described previously and are inserted into a single location, that is, the olecranon. Therefore, in MRI, this characteristic takes on a bipartite appearance.[13-15]

An accurate understanding of the distal anatomy of the triceps tendon insertion is clinically important in surgical planning for procedures such as reduction of displaced fractures of the distal portion of the posterior surface of the humerus, olecranon osteotomies and repairs of partially or completely torn triceps tendons. Furthermore, it is important to know that the most common traumatic injuries associated with triceps tendon, is a fracture of the radial head, probably due to the same injury mechanism (fall with the arm extended). Other possible associated injuries are medial collateral ligament tears or laxity, compression of the ulnar or radial nerve as well as fractures of the proximal humerus. Immediate surgical repair of an acute tear (less than six weeks) is generally recommended with reinsertion of the triceps tendon using suture anchors. When treatment is delayed, the degenerated and fragile appearance of the stump can prevent reinsertion of the tendon. [14-16]

In this study, a detailed assessment of the triceps brachii tendon was carried out retrospectively in in-vivo MRI studies, resulting in an assertive assessment in relation to some studies that used cadavers to perform elbow studies.[13]

It is important to remember that the main difference of this study in relation to others already carried out using MRI is that this study was evaluated by examiners as well as evaluation of retrospective in vivo exams, showing a single insertion of the triceps brachii in Figures 3 and 4. Therefore, there are no changes in relation to the post-mortem. For example, with a reduction in the level of hydration of the tendon and cadaver stiffness as well as possible previous degenerative changes, these changes that were excluded from the study, could compromise the evaluation.

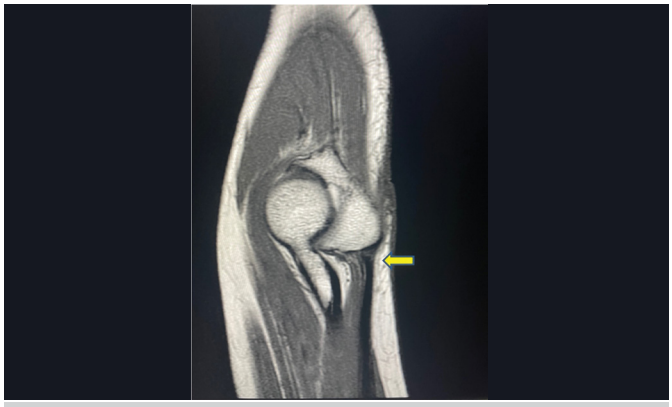


Figure 3. T1 Sequence MRI sagittal scan showing of the distal triceps insertion in the olecranon (yellow arrow).

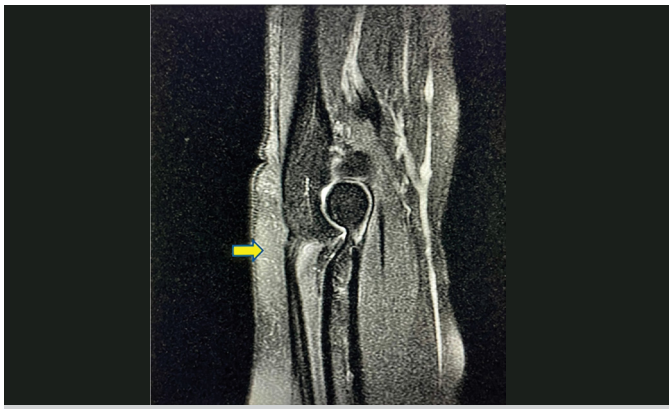


Figure 4. T2 Sequence MRI sagittal scan showing the distal triceps tendon insertion in the olecranon (yellow arrow).

Although Belentani et al.,[13] using a different MRI protocol, describes the distal insertion of the triceps tendon as a bipartite appearance. However, histological analysis in this study provided definitive evidence of a single distal insertion.[13] This same observation is also reported by Hayter and Adler.[6-8] t was identified that there is a tendon of the distal medial head that joins the lateral and long tendons (sets). In the future with better and enhanced MRI equipment, this tendon could be more visible since, currently; it is too small to authenticate its existence. However, even with this aspect the insertion was identified as unique.

In MRI studies the main tendon of the medial head was not identified. Barco et all[15] indicates that the medial head tendon is inconsistent or too small to be measured. Another possible explanation for the different MRI and histology results analyzed in this study[14] is the histological sections. They showed the tip of the insertion where the central tendon is very close to the medial head tendon and, probably, there could be fusion of the fibers of the three heads. The rolled edge represents a deep medial thickening of the central tendon that receives contributions from the medial head.[12] However, a survey of the integrity of all insertion sites of the medial head also identifies a single insertion in the olecranon in the deep muscle component.[14-17,18]

A distal pre-tricipital space found in anatomical dissections, probably a split, demonstrates a bipartite appearance, as previously described in the literature.[11] As in the study by Akamatsu et. all,[19] some MRI scans spotted the existence of the pre-tricipital space. However, it is also not clearly identified by observing radiologists.[19]

According to a study by Kholinne et. all,[21] this current study also has limitations. However, this retrospective study, with a greater number of patients in vivo, provides a greater range of analysis because the MRI studies were carried out on living patients and not associated with cadavers. Therefore, the evaluation is not hampered by reduced post-mortem hydration, as well as cadaver stiffness, which we believe may present less accurate results.[20-21]

The area is not measured in the MRI study because the traditional measurement at tendon insertion is calculated by the greatest length x width. Thus, one can accurately underestimate the 3D structure of its actual insertion area.

The lack of histological evaluation or dissection is due to the impossibility of performing histological studies or dissection on living patients. Despite the previously mentioned limitations, this study provided us with a clear view that there is a single insertion of the triceps brachii tendons into the olecranon, even taking into consideration that the tendons have a different form of insertion.

Therefore, this information provided in this study can help to clarify the diagnosis of partial ruptures of the distal triceps brachii tendon, helping to provide a reliable satisfactory result for the patient.

CONCLUSION

The detailed study of the insertion region of the triceps brachii tendon using MRI identifies and confirms that the triceps brachii tendons have a single insertion into the olecranon.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. NJR: writing, analysis of magnetic resonance images, and manuscript preparation. FNAA: data analysis, assistance, and writing. COP: article review and intellectual concept of the article.

DATA AVAILABILITY DECLARATION

The underlying contents of the research text are contained in the manuscript.

REFERENCES

1. Testut L, Latarjet A. Tratado de Anatomia Humana. 9th ed., reprint. Salvat Editores S.A.; 1979. p. 1019-23.
2. Gray H. Gray's Anatomy. The Anatomical Basis of Medicine and Surgery. 38th ed. Churchill Livingstone; 1995.
3. Gray H. Gray's Anatomy. 40th ed., reprint. London: Churchill Livingstone; 2008.
4. Sampath SC, Sampath SC, Bredella MA. Magnetic resonance imaging of the elbow: A Structured Approach. Sports Health. 2012;5(1):34-9. doi: 10.1177/1941738112467941.
5. Bucknor M, Stevens K, Steinbach L. Elbow imaging in sport: Sports imaging series. Radiology. 2016;279(1):12-28. doi: 10.1148/radiol.2016150501.
6. Athwal GS, McGill RJ, Rispoli DM. Isolated avulsion of the medial head of the triceps tendon: an anatomic study and arthroscopic repair in 2 cases. Arthroscopy. 2009;25(9):983-8. doi: 10.1016/j.arthro.2009.02.020.
7. Chan APH, Lo CK, Lam HY, Fung KY. Unusual traumatic triceps tendon rupture: a word of caution. Hong Kong Med J. 2009;15(4):294-6.
8. Hayter CL, Adler RS. Injuries of the elbow and the current treatment of tendon disease. AJR Am J Roentgenol. 2012;199(3):546-57. doi: 10.2214/AJR.11.8325.
9. Sharma SC, Singh R, Goel T, Singh H. Missed diagnosis of the triceps tendon rupture: a case report and review of literature. J Orthop Surg. 2005;13(3):307-9. doi: 10.1177/230949900501300317.
10. Levy M. Repair of triceps tendon avulsions or ruptures. J Bone Joint Surg Br. 1987;69(1):115. doi: 10.1302/0301-620X.69B1.3546325.
11. Keener JD, Chafik D, Kim HM. Insertional anatomy of the triceps brachii tendon. J Shoulder Elbow Surg. 2010;19(3):399-405. doi: 10.1016/j.jse.2009.10.008.
12. Madsen M, Marx R, Millett P, Rodeo SA, Sperling JW, Warren RF. Surgical anatomy of the triceps brachii tendon anatomical study and clinical correlation. Am J Sports Med. 2006;34(11):1839-43. doi: 10.1177/0363546506288752.
13. Belentani C, Pastore D, Wangwinyuvirat M, Dirim B, Trudell DJ, Haghighi P, et al. Triceps brachii tendon: anatomic-MR imaging study in cadavers with histologic correlation. Skeletal Radiol. 2009;38(2):171-5. doi: 10.1007/s00256-008-0605-6.
14. Negrão JR, Mogami R, Ruiz FAR, Wagner FV, Haghighi P, Ward SR, et al. Distal insertions anatomy of the triceps brachii muscle: mri assessment in cadaver specimens employing histologic correlation and play-doh® models of the anatomic findings. Skeletal Radiol. 2020;49(7):1057-67. doi: 10.1007/s00256-020-03382-3.
15. Morrey BF. Anatomy of the elbow joint and its disorders. 2nd ed. Philadelphia W.B. Saunders; 1993. p. 16-52.
16. Jacobson JA, Lebson PJL, Jeffers AW, Fessel D, Hayes C. Ulnar nerve dislocation snapping triceps syndrome: diagnosis with dynamic sonography-report of three cases. Radiology. 2001;220(3):601-5. doi: 10.1148/radiol.2202001723.
17. Evans RC. O Cotovelo. In: Exame físico ortopédico ilustrado. 2nd ed. São Paulo: Manole; 2003. p. 276-324.
18. Keener JD, Chafik D, Kim M, Galatz LM, Yamaguchi K. Insertional anatomy of the triceps brachii tendon. J Shoulder Elbow Surg. 2010;19(3):399-405. doi: 10.1016/j.jse.2009.10.008.
19. Akamatsu FE, Negrão JR, Rodrigues MB, Itezerote AM, Saleh SO, Hojaij F, et al. Is there something new regarding tríceps brachii muscle insertion?. Acta Cir Bras. 2020;35(10):e202001007. doi: 10.1590/s0102-865020200100000007.
20. Celli A. Triceps tendon rupture: knowledge acquired from the anatomy to the surgical repair. Musculoskelet Surg. 2015;99(Suppl1):S57-66. doi: 10.1007/s12306-015-0359-y.
21. Kholinne E, Kwak JM, Heo Y, Hwang SJ. The anatomic – magnetic resonance imaging study of distal triceps brachii tendon. J Orthop Surg (Hong Kong). 2022;30(3):10225536221122262. doi: 10.1177/10225536221122262.

DOI: http://dx.doi.org/10.1590/1413-785220263404e300989

ANTHROPOMETRIC CORRELATION OF HAMSTRING AND PERONEUS LONGUS GRAFTS IN ACL RECONSTRUCTION

CORRELAÇÃO ANTROPOMÉTRICA DOS ENXERTOS FLEXORES E DO FIBULAR LONGO NA RECONSTRUÇÃO DO LCA

Izabela Feres de Oliveira[1], *Matheus Kuster Ronconi*[1], *Gabriel Velasque dos Santos Midão*[1], *Fabrício Nascimento de Almeida*[1], *Gustavo Dalla Bernardina de Almeida*[1], *Saulo Gomes de Oliveira*[1]

1. Hospital Santa Casa de Misericordia de Vitoria, Vitoria, ES, Brazil.

ABSTRACT

This study aimed to analyze the association between anthropometric data of patients with anterior cruciate ligament injuries and the characteristics of the grafts used in ligament reconstruction. This analytical study involved 30 patients randomly assigned to two groups: Group I, flexor tendon grafts (semitendinosus and gracilis), and Group II, peroneus longus tendon grafts. Graft length and diameter were measured, as well as preoperative anthropometric variables such as age, height, and weight. Student's t-test, chi-square test, and Pearson's correlation test were used. The results showed significantly greater length for the flexor tendon grafts (28.8 ± 2.66 vs. 25.8 ± 3.19 cm), while the peroneus longus tendon exhibited a greater final diameter (9.33 ± 0.82 vs. 8.47 ± 0.52 cm), both with $p < 0.05$. A negative correlation was observed between age and the length of the peroneus longus graft, and a positive correlation between height and the same parameter. Anthropometric data were not shown to be reliable predictors for the flexor grafts, although younger and taller patients were more likely to have a greater length in the peroneus longus graft. ***Level of Evidence I; Prognostic study – investigation of the effect of a patient's characteristics on disease outcome.***

Keywords: Anthropometry; Anterior Cruciate Ligament Injuries; Transplants; Tendons.

RESUMO

O estudo objetivou analisar a associação entre os dados antropométricos de pacientes com lesão do ligamento cruzado anterior e as características dos enxertos empregados na reconstrução ligamentar. Trata-se de um estudo analítico com 30 pacientes, distribuídos aleatoriamente em dois grupos: Grupo I, enxerto dos tendões flexores (semitendíneo e grácil), e Grupo II, enxerto do tendão fibular longo. Foram mensurados o comprimento e o diâmetro dos enxertos, além de variáveis antropométricas no pré-operatório, como idade, estatura e peso. Utilizou os testes t de Student, Qui-quadrado e Correlação de Pearson. Os resultados mostraram comprimento significativamente maior para os enxertos dos tendões flexores (28,8 ± 2,66 vs. 25,8 ± 3,19 cm,), enquanto o tendão fibular longo exibiram diâmetro final superior (9,33 ± 0,82 vs 8,47 ± 0,52 cm), ambos com $p < 0,05$. Observou-se correlação negativa entre idade e comprimento do enxerto fibular longo, e correlação positiva entre estatura e o mesmo parâmetro. Os dados antropométricos não se mostraram preditores confiáveis para os enxertos flexores, embora pacientes mais jovens e mais altos tenham maior probabilidade de comprimento superior no enxerto fibular longo. ***Nível de Evidência I; Estudo prognóstico – investigação do efeito de característica de um paciente sob o desfecho da doença.***

Descritores: *Antropometria; Lesões do Ligamento Cruzado Anterior; Transplantes; Tendões.*

Citation: Oliveira IF, Ronconi MK, Midão GVS, Almeida FN, Almeida GDB, Oliveira SG. Anthropometric correlation of hamstring and peroneus longus grafts in ACL reconstructio. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 6. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

The anterior cruciate ligament (ACL) originates from the posterior face of the lateral femoral condyle and inserts laterally and anteriorly to the medial tibial spine. The ACL has an average intra-articular length of 38 mm and an average diameter of 11 mm,[1] with its primary function being the anterior stabilization of the tibia and restriction of internal rotation of the knee.[2]

ACL rupture is the most common ligament injury of the knee.[3] Among the recommended treatments, surgical method is the primary choice, with intra-articular reconstruction via arthroscopy being the most common.[4] According to the Ministry of Health in Brazil, the incidence of ACL reconstruction procedures is 3.49 cases per 100,000 people per year, highlighting the importance of this treatment in the Brazilian population within the Unified Health System (SUS).[5]

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Hospital Santa Casa de Misericordia de Vitoria, Vitoria, Espirito Santo, Brazil.
Correspondence: Izabela Feres de Oliveira. Hospital Santa Casa de Misericordia de Vitoria, Vitoria, ES, Brazil. 29025-023. izabela.feres@gmail.com

Article received on 09/16/2025 approved on 09/24/2025.
Handling Editor: Fabio Janson Angelini

During ACL reconstruction, the choice of the appropriate graft is crucial for the success of the procedure. Autologous grafts, such as those from the quadriceps tendons, patellar tendon, flexors (semitendinosus and gracilis), and peroneus longus, are commonly used due to their availability and biocompatibility.[6] Studies show that grafts with diameters less than 7 mm present a higher risk of failure and recurrence,[7] emphasizing the importance of adequate size to ensure postoperative durability. Each type of graft has its particularities regarding length, diameter, and presents specific advantages and disadvantages concerning postoperative functional impairment, such as muscle recovery of the donor area, postoperative pain, and risk of long-term complications.[8]

It is known that the long fibular tendon graft has high resistance, safety, and lower morbidity; however, few studies demonstrate its surgical disadvantages, unlike the hamstrings, which are widely studied, already showing loss of strength and a long incorporation time (ligamentization).[9]

In light of the above, this study aims to compare the autologous grafts of the flexors (semitendinosus + gracilis) (STG) to that of the long fibular (FL), correlating these data with the anthropometric characteristics of the patients, collected preoperatively. In order to analyze whether there is a statistical difference between the characteristics of the grafts, Student's *t* tests (simple and paired) and Chi-Square tests were used to evaluate differences between groups and associations. Additionally, correlations between anthropometric data and graft characteristics were analyzed using Pearson correlation, whose values range from -1 to 1, indicating a direct or inverse relationship.

MATERIALS AND METHODS

This is an observational, cross-sectional, analytical correlation study involving patients with complete ACL rupture, scheduled for RACL reconstruction at the knee outpatient clinic of a philanthropic hospital, between June and December 2024. The research was initiated after approval from the Ethics Committee with CAAE number 82144124.2.0000.5065.

Thirty patients were included, randomly divided by the Research Randomizer site into two groups of 15 individuals, according to the type of autologous tendon harvested for RACL reconstruction, being: Group I, flexor graft (semitendinosus + gracilis tendons); and Group II, long fibular tendon.

The collection of anthropometric information was conducted through a questionnaire during the preoperative consultation, after they agreed and signed the Informed Consent Form (ICF).

The patient's weight was measured using a digital scale with a precision of one decimal place. Height was measured with a non-elastic tape measure with the patient in an orthostatic position with bipodal support.[10] Other anthropometric data were obtained through a questionnaire.

In Group I, the semitendinosus and gracilis tendons were harvested through a 2–3 cm incision on the anteromedial face of the knee, at the insertion of the Pes Anserinus, about 4–5 cm below the medial joint line,[11] using Stripper instruments. Subsequently, they were subjected to the removal of muscle fragments and prepared on the surgical table with the aid of forceps and high-strength sutures (Ethibond No. 2), where their longitudinal edge was sutured to keep the edges of the semitendinosus and gracilis tendons together, using Krakow suture (Figure 1). After this preparation, their length was measured with a sterile millimeter ruler, obtaining the final length of the graft. Finally, the graft was folded according to the plan, achieving a minimum final length of 9 cm, with its diameter measured using a standardized gauge (Figure 2).

In Group II, the graft was harvested through an incision on the lateral face of the ankle, posterior to the lateral malleolus, with the aid of the Stripper. The distal portion was preserved and sutured to the short fibular tendon (tenodesis) to maintain part of the muscular function of the long fibular (Figure 3).[12] The same measurements were taken in Group I after preparation.

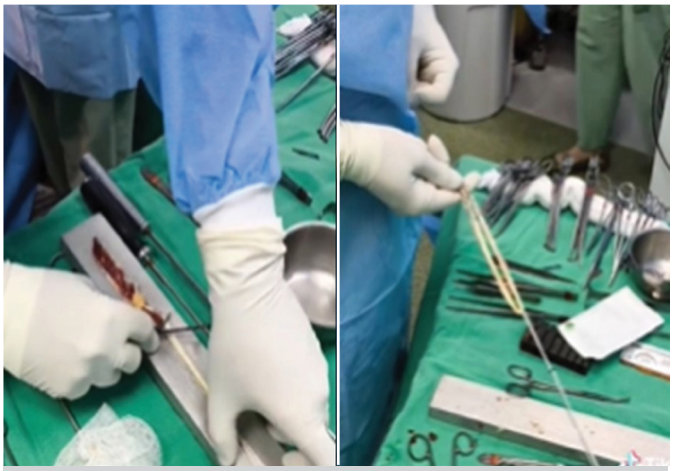


Figure 1. Preparation of the collected graft.

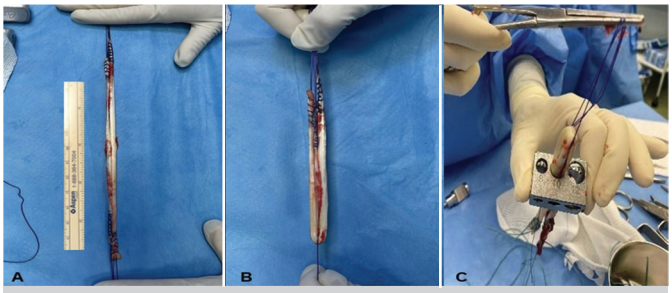


Figure 2. A- Measurement of the graft length after preparation, using calibrated surgical instruments. B- Graft folded after measuring the length. C- Measurement of the graft diameter with a standardized gauge.

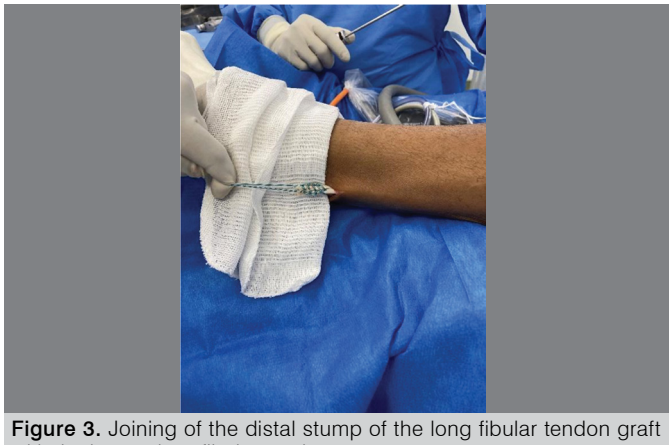


Figure 3. Joining of the distal stump of the long fibular tendon graft with the intact short fibular tendon.

The data obtained from the anthropometric measurements and grafts were transferred to a table in the software Microsoft Office/Excel 2011 (Redmond, Washington, USA).

The statistical methodology used Student's t-tests (simple and paired) to assess differences between groups, and the Chi-Square test to identify associations. The correlations between anthropometric data and graft characteristics were analyzed using Pearson's correlation, whose values range from -1 to 1, indicating a direct or inverse relationship.

A confidence interval for the mean and p-value was used, with a significance level of 0.05. The data from the grafts and anthropometric correlations were analyzed using parametric or non-parametric tests, depending on the distribution. The statistical analysis was performed using the software SPSS Statistics 28.0.1. P values < 0.05 indicated statistical significance.[13]

RESULTS

The Table 1 shows a comparison of the anthropometric factors of Groups I (Flexor) and II (Fibular) regarding height, age, and weight, where the values ranged from 166-195 cm, 17-52 years, and 68-97 kg (Flexor), and 153-185 cm, 22-66 years, and 60-117 kg (Fibular). In general, the results demonstrated an absence of statistical difference, indicating homogeneity between the groups.
The comparative analysis of the mean lengths of the grafts showed a statistically significant difference, with a greater mean length in Group I (Flexors), 28.80 cm (±2.66), compared to Group II (Long Fibular), 25.80 cm (±3.19), with p=0.009. Additionally, there was a difference in the mean diameters, greater in Group II, Long Fibular graft, 9.33 mm (±0.82) compared to Group I, Flexor graft, 8.47 (±0.52) with (p 0.002) (Table 2 and Figure 4)
The correlation between anthropometric measurements (age, height, and weight) with the length and diameter of the grafts was analyzed in both groups (Table 3). Age showed a negative correlation with the length of the Long Fibular graft, (r= -0.622), with (p=0.013), indicating an inversely proportional variation, where a lower age corresponds to a greater length of the obtained long fibular graft. Height showed a positive correlation with the length of the Long Fibular graft, (r= 0.736), with (p=0.002), indicating a directly proportional variation, where a greater height of the patient corresponds to a greater length of the obtained long fibular graft. These variables did not show a statistically significant correlation with the flexor tendon graft, nor did weight in either of the two groups. (Figures 5-8).

Table 1. Comparison of the Anthropometric Factors of Groups I (Flexor) and II (Fibular).

		Mean	Median	Standard Deviation	CV	Min	Max	N	IC	P-value
Age	Flexor	30.3	25	12.5	41%	17	52	15	6.3	0.189
	Fibular	37.1	38	14.8	40%	22	66	15	7.5	
Height	Flexor	176.0	175	7.4	4%	166	195	15	3.7	0.065
	Fibular	170.3	170	8.7	5%	153	185	15	4.4	
Weight	Flexor	80.5	77	9.1	11%	68	97	15	4.6	0.354
	Fibular	85.0	80	16.3	19%	60	117	15	8.2	

Table 2. Comparison of groups I and II of the average length and diameter of the completed grafts.

		Average	Median	Standard Deviation	CV	Min	Max	N	IC	P-value
Length of the Completed Graft	Group I	28.80	28.5	2.66	9%	24	33	15	1.31	0.009
	Group II	25.80	26	3.19	12%	20	31	15	1.61	
Diameter of the Completed Graft	Group I	8.47	8	0.52	6%	8	9	15	0.26	0.002
	Group II	9.33	9	0.82	9%	8	11	15	0.41	

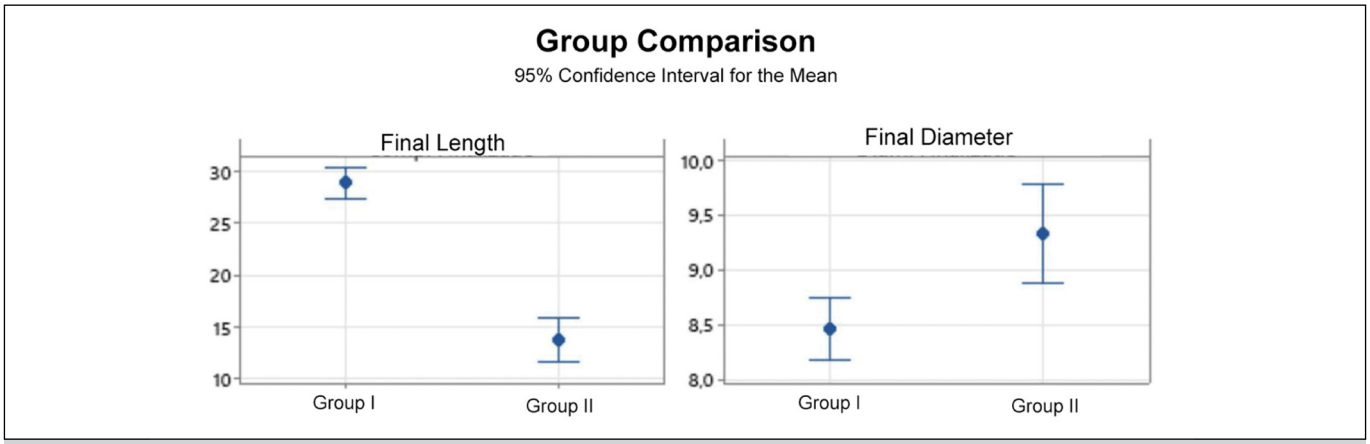


Figure 4. Comparison of length (cm) and diameter (cm) of the finished grafts by group (I and II).

Table 3. Correlation of anthropometric factors with tendon measurements by group.

		Age		Height		Weight	
		Corr (r)	P-value	Corr (r)	P-value	Corr (r)	P-value
Group I	Completed Length	-0.248	0.372	0.404	0.135	0.354	0.196
	Completed Diameter	-0.247	0.375	-0.056	0.843	0.240	0.389
Group II	Length of the Completed Graft	-0.518	0.048	0.475	0.073	-0.047	0.868
	Diameter of the Completed Graft	0.164	0.560	0.265	0.341	0.258	0.353

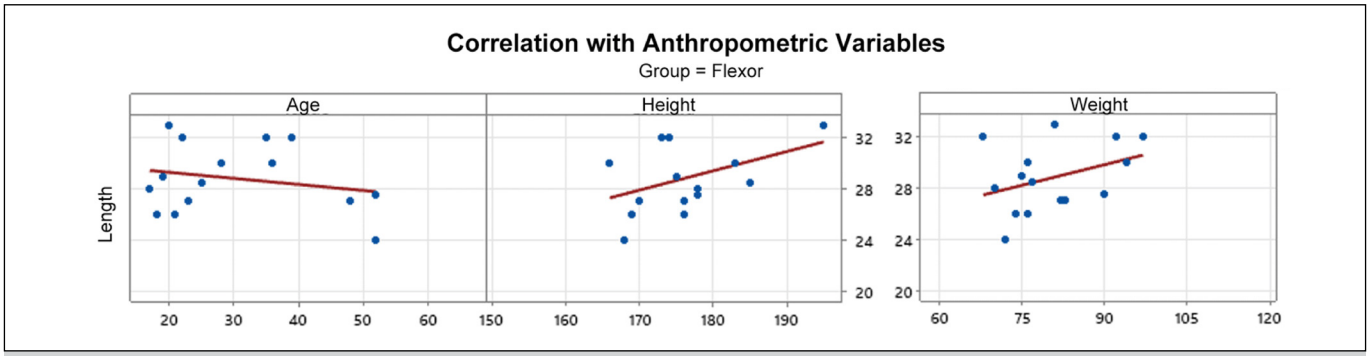


Figure 5. Correlation between the anthropometric measurements of age, height, and weight with the parameter of the finalized graft length, Group I, Flexor Tendon.

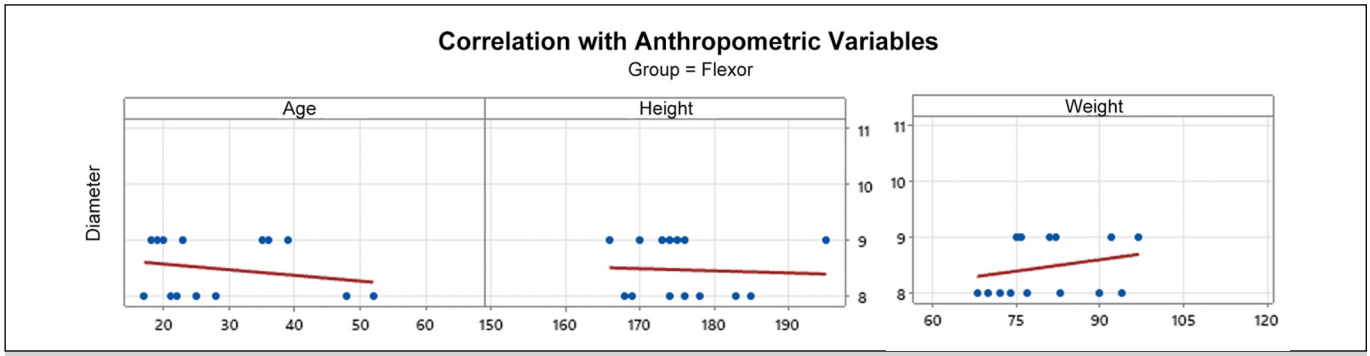


Figure 6. Correlation between the anthropometric measurements of age, height, and weight with the parameter of the finalized graft diameter, Group I, Flexor Tendon.

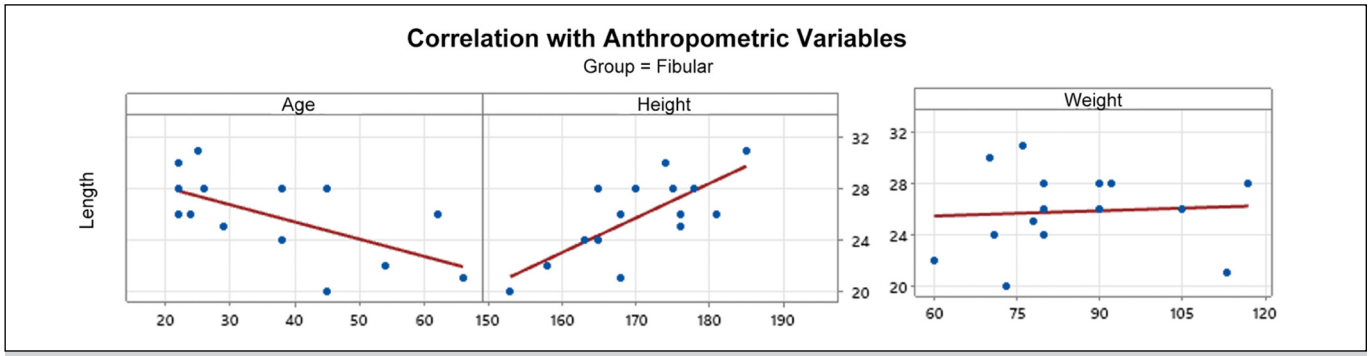


Figure 7. Correlation between the anthropometric measurements of age, height, and weight with the parameter of the finalized graft length, Group II, Long Fibular Tendon.

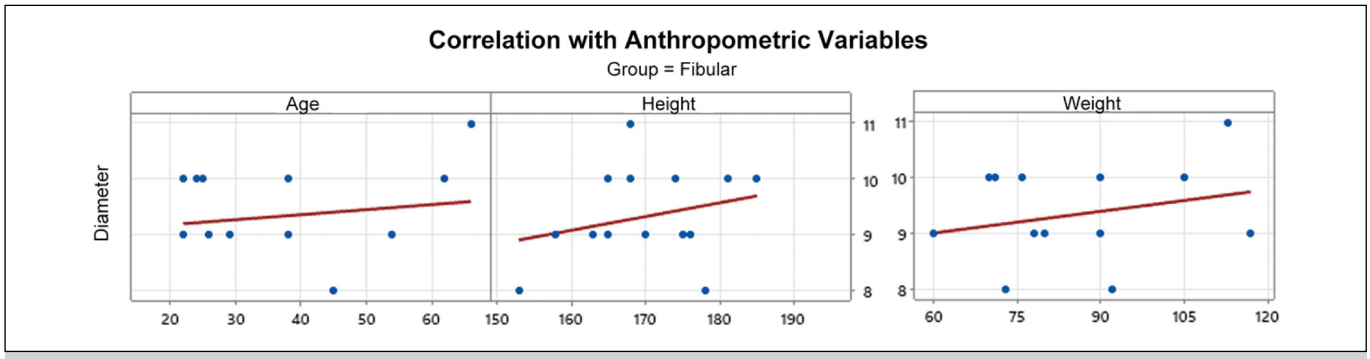


Figure 8. Correlation between the anthropometric measurements of age, height, and weight with the parameter of the finalized graft diameter, Group II, Long Fibular Tendon.

DISCUSSION

When comparing our findings with the literature, we found an average length of the flexor tendons after preparation, before being folded, of (28.80 ± 2.66 cm), which was considerably higher than the flexor graft reported in other studies. Wan et al[14] found average lengths of flexor grafts close to 25 cm.

Regarding the length of the Long Fibular graft, Khan et al.,[15] report an average length of 321.4 ± 26.7 mm, a value higher than that found in our study (25.80 ± 3.19 cm), however, a sufficient length, as stated by Janssen et al.,[16] who indicated minimum lengths of 21 cm sufficient for complex anatomical ligament reconstructions of the ACL.

Goyal et al.,[17] highlighted the minimum graft length for ACL reconstruction, finalized after the fold, of 8 cm in length. Our case series achieved a graft length that allows for folds in both groups, without compromising the useful length, ensuring adequate filling of the bone tunnels and stability of the fixation.

Similarly, the length of the found flexor tendons allows for use in combined or multiligament reconstructions, where several distinct tunnels and greater tissue demand are necessary, especially in the group of Flexors that were statistically larger than the Long Fibular tendon, whose use as a graft in these surgeries requires caution due to the lack of adequate length in some situations.[18]

The long fibular graft had a statistically greater diameter than the flexor in our case series: greater in relation to Group I (Flexors): 8.47±0.52 vs. Group II, (Long Fibular): 9.33 mm±0.82 (p= 0.002). Gupta et al. (2018), evaluating 156 patients undergoing ACL reconstruction with long fibular tendon graft, found an average diameter of 8.3 mm, having a significant correlation with height, weight, and injury duration, allowing for the development of a predictive equation for graft caliber.[19] Our results demonstrated an average diameter of the Long Fibular graft superior to that reported by Gupta et al., which may be justified by distinct anthropometric and ethnic factors in each sample.

We also observed a significant negative correlation between the length of the long fibular graft and the age of the patients ($r = -0.518$; $p = 0.048$), indicating that younger individuals tend to have larger caliber grafts. Thus, we reinforce the likely biomechanical advantage of the long fibular as a safe and effective graft, demonstrating it to be an excellent graft option in this group of patients, given that grafts ≥ 8 mm present a lower risk of failure.[18] However, more clinical and biomechanical studies need to confirm this thesis, reinforcing whether there is superiority in ACL reconstructions with Long Fibular graft in younger patients, given the larger average diameter in this population.

Another finding demonstrates a positive correlation between height and final length of the long fibular graft, being directly proportional to the greater height of the patient with the longer fibular graft. We did not find data or studies evaluating these characteristics in the medical literature, possibly due to the recent interest in the long fibular graft and few published works on the subject.

Thus, considering that there was a statistical correlation between the anthropometric data and the measurements of the grafts, and that the grafts of the flexor tendons are longer and those of the long fibular tendon have a greater diameter, it is possible to guide surgical planning according to the need: opting for the flexors when greater length is required, and for the long fibular when there is a need for a graft with greater diameter, with possible advantages of stiffness and resistance.

CONCLUSION

The results of this study demonstrated that the grafts of the flexor tendons have significantly greater length, while the long fibular tendon has a superior final diameter, both with statistical significance. It was also observed that age presented a negative correlation and height a positive correlation with the length of the long fibular graft, while the other anthropometric variables did not show reliable predictors. These findings reinforce the importance of considering the specific characteristics of each graft in the choice of the reconstructive technique for the anterior cruciate ligament, indicating that the long fibular tendon may represent a valid alternative when seeking greater diameter, especially in younger and taller patients.

ACKNOWLEDGMENTS

I would like to thank the patients who kindly agreed to participate in this study. Without your trust and willingness, this work would not be possible. And to the healthcare professionals involved directly or indirectly in the process, who contributed to the completion of the research. The competence and commitment of each individual were fundamental to the success of this study.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. AFN, AGDB, and OSG: performing surgeries and reviewing the article; MGVS: data analysis; RMK and OIF: writing, article review, and the intellectual concept of the article.

DATA AVAILABILITY DECLARATION

The underlying content of the research text is contained in the manuscript.

REFERENCES

1. Kupczik F, Schiavon MEG, Sbrissia B, Fávaro RC, Valério R. Enxerto ideal para ligamento cruzado anterior: correlação em ressonância magnética entre LCA, isquiotibiais, tendão patelar e tendão quadríceps. Rev Bras Ortop. 2013;48(5):441-447. doi:10.1016/j.rbo.2012.11.003.
2. Jobe CM. Rotator cuff disorders: anatomy, function, pathogenesis, and natural history. In: Norris TR, editor. Orthopaedic Knowledge Update: Shoulder and Elbow. Rosemont (IL): American Academy of Orthopaedic Surgeons; 2002. p. 143-154.
3. Magnussen RA, Lawrence JTR, West RL, Toth AP, Taylor DC, Garrett WE. Graft size and patient age are predictors of early revision after anterior cruciate ligament reconstruction with hamstring autograft. Arthroscopy. 2012;28(4):526-531. doi:10.1016/j.arthro.2011.11.024.
4. Odensten M, Gillquist J. Functional anatomy of the anterior cruciate ligament and a rationale for reconstruction. J Bone Joint Surg Am. 1985;67(2):257-262.
5. Lopes TJA, Simic M, Pappas E. Epidemiology of anterior cruciate ligament reconstruction in Brazil's public health system. Rev Bras Med Esporte. 2016;22(4):297-301. doi:10.1590/1517-869220162204159074.
6. Rezende FC. Reconstrução combinada intra e extra-articular versus reconstrução intra-articular isolada do ligamento cruzado anterior: revisão sistemática [dissertação]. São Paulo: Universidade Federal de São Paulo; 2016.
7. Santos JRT, Carvalho EPS, Silva MLG, Bonfim MAL, Santos LRT, Carvalho MEIM. Epidemiologic profile of patients submitted to reconstruction of the anterior cruciate ligament in a referral public hospital in Teresina, PI, Brazil. Fisioter Bras. 2015;16(2):84-88.
8. Pagnani MJ, Warner JJP, O'Brien SJ, Warren RF. Anatomic considerations in harvesting the semitendinosus and gracilis tendons and a technique of harvest. Am J Sports Med. 1993;21(4):565-571. doi:10.1177/036354659302100414.
9. Melo AMT, Neto HSB, Jesus PHS, Menezes ICG, Lopes GPG, Moreno VPD, et al.

Reconstrução do ligamento cruzado anterior: uma revisão bibliográfica acerca dos enxertos utilizados. Rev FT. 2023;27(121). doi:10.5281/zenodo.7829268.
10. Petroski EL. Padronização de medidas antropométricas e avaliação da composição corporal. Porto Alegre: CIPA; 2007.
11. Vinagre G, Kennedy NI, Chahla J, Cinque ME, Hussain ZB, Olesen ML, et al. Hamstring graft preparation techniques for anterior cruciate ligament reconstruction. Arthrosc Tech. 2017;6(6):e2079-e2084. doi:10.1016/j.eats.2017.08.031.
12. Arora M, Shukla T. Peroneus longus graft harvest: a technique note. Indian J Orthop. 2023;57(4):611-616. doi:10.1007/s43465-023-00847-0.
13. Spiegel MR. Estatística. 3a ed. São Paulo: Makron Books; 1993.
14. Wan C, Zhou Y, Li H, Zhang H, Li Y, Wang J, et al. Comparison of clinical outcomes using 4-strand versus 5-strand hamstring tendon autograft for anterior cruciate ligament reconstruction. Asia Pac J Sports Med Arthrosc Rehabil Technol. 2022;28:35-40. doi:10.1016/j.asmart.2022.06.002.
15. Khan MA, Umar M, Shah NH, Nadeem RD. Prediction of peroneus longus graft diameter for anterior cruciate ligament reconstruction. Chin J Traumatol. 2021;24(6):320-324. doi:10.1016/j.cjtee.2021.06.007.
16. Janssen RPA, van der Velden MJF, Pasmans HLT, Sala HAGM. Length of hamstring tendons in Caucasians: an anthropometric study in patients undergoing ACL reconstruction. Knee Surg Sports Traumatol Arthrosc. 2015;23(3):685-690. doi:10.1007/s00167-013-2452-2.
17. Goyal S, Matias N, Pandey V, Acharya K. Minimum essential autograft dimensions for single-bundle anterior cruciate ligament reconstruction using quadrupled hamstring graft: a retrospective analysis. Knee Surg Relat Res. 2019;31(1):2-8. doi:10.1186/s43019-019-0002-6.
18. Gupta H, Doddamani RS, Dwidmuthe S, Goel A, Verma R. Predicting the graft diameter of the peroneus longus tendon for anterior cruciate ligament reconstruction: a pilot study. J Orthop. 2018;15(2):711-715. doi:10.1016/j.jor.2018.05.036.
19. Silva A, Helito CP, Bonadio MB, Demange MK, Gobbi RG, Pécora JR, et al. Combined anterior cruciate ligament and anterolateral ligament reconstruction: minimum graft length requirements. Rev Bras Ortop. 2023;58(1):77-83. doi:10.1055/s-0042-1758711.

DOI: http://dx.doi.org/10.1590/1413-785220263404e303555

BACTERIAL COLONIZATION IN ORTHOPEDIC SURGICAL TOURNIQUETS: A MULTICENTER STUDY

COLONIZAÇÃO BACTERIANA EM TORNIQUETES CIRÚRGICOS ORTOPÉDICOS: ESTUDO MULTICÊNTRICO

Eduardo Frois Temponi[1], *Luiz Lorentz Salzmann Lamego*[1], *Eduardo Louzada da Costa*[2], *Glauco Almeida Passos*[3,4], *Tiago Soares Baumfeld*[4,5], *Leandro da Silva Marinho*[6], *Túlio Vinícius de Oliveira Campos*[7], *Raquel Bandeira da Silva*[8], *Vitor Rodrigues Miranda*[1]

1. Pontifícia Universidade Catolica de Minas Gerais, Hospital Madre Teresa, Belo Horizonte, MG, Brazil.
2. Hospital da Unimed BH, Belo Horizonte, MG, Brazil.
3. Hospital Evangelico de Belo Horizonte, Belo Horizonte, MG, Brazil.
4. Hospital Felicio Rocho, Belo Horizonte, MG, Brazil.
5. Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil.
6. Hospital da Unimed, Hospital da Baleia, Belo Horizonte, MG, Brazil.
7. Universidade Federal de Minas Gerais, Hospital Risoleta Tolentino Neves, Belo Horizonte, MG, Brazil.
8. Hospital Universitario Ciencias Medicas, Faculdade de Medicina Ciencias Medicas, Belo Horizonte, Minas Gerais, Brasil.

ABSTRACT

Objective: To evaluate the prevalence and microbiological profile of contamination in reusable tourniquets in Brazilian hospitals. Methods: A multicenter study conducted in six hospitals. Swabs were collected from 54 tourniquets immediately after surgical use and before disinfection, covering an estimated area of 10 cm². The samples were cultured and identified using automated methods. The microbial load was described by median and interquartile range (IQR), and comparisons between public and private hospitals were performed using Fisher's exact test and Mann-Whitney test. Results: The prevalence of contamination was 70.4% (38/54). The median overall microbial load was 101 CFU per device (IQR: 0–153), corresponding to approximately 10.1 CFU/cm². The predominant microorganisms were coagulase-negative Staphylococcus (48.1%) and Staphylococcus aureus (18.5%), with isolation of Pseudomonas aeruginosa, Bacillus sp., and Candida sp. The contamination rate was 78.6% in public hospitals and 61.5% in private hospitals (p=0.081), with no statistically significant difference in the median bacterial load between the institutions (p=0.412). Conclusion: There is a high prevalence of contamination by clinically relevant pathogens in reusable tourniquets, regardless of the type of hospital. The results indicate systemic failures in reprocessing and suggest the need for high-level disinfection protocols or the adoption of disposable sterile devices to mitigate the risk of cross-contamination. ***Level of evidence III; multicenter cross-sectional study of microbiological prevalence.***

Keywords: Tourniquets; Cross Infection; Orthopedic Procedures; Bacterial Infections; Brazil

RESUMO

Objetivo: Avaliar a prevalência e o perfil microbiológico da contaminação em torniquetes reutilizáveis em hospitais brasileiros. Métodos: Estudo multicêntrico conduzido em seis hospitais. Foram coletados swabs de 54 torniquetes imediatamente após o uso cirúrgico e antes da desinfecção, abrangendo uma área estimada de 10 cm². As amostras foram cultivadas e identificadas por métodos automatizados. A carga bacteriana foi descrita por mediana e intervalo interquartil (IIQ), e as comparações entre hospitais públicos e privados foram realizadas pelos testes exato de Fisher e de Mann-Whitney. Resultados: A prevalência de contaminação foi de 70,4% (38/54). A mediana da carga microbiana geral foi de 101 UFC por dispositivo (IIQ: 0–153), correspondendo a aproximadamente 10,1 UFC/cm². Os microrganismos predominantes foram Staphylococcus coagulase-negativo (48,1%) e Staphylococcus aureus (18,5%), com isolamento de Pseudomonas aeruginosa, Bacillus sp. e Candida sp. A taxa de contaminação foi de 78,6% nos hospitais públicos e de 61,5% nos hospitais privados (p=0,081), sem diferença estatisticamente significativa na mediana da carga bacteriana entre as instituições (p=0,412). Conclusão: Há alta prevalência de contaminação por patógenos clinicamente relevantes em torniquetes reutilizáveis, independentemente do tipo de hospital. Os resultados indicam falhas sistêmicas no reprocessamento e sugerem a necessidade de protocolos de desinfecção de alto nível ou a adoção de dispositivos estéreis descartáveis para mitigar o risco de contaminação cruzada. ***Nível de evidência III; estudo transversal multicêntrico de prevalência microbiológica.***

Descritores: *Torniquetes; Infecção Cruzada; Procedimentos Ortopédicos; Infecções Bacterianas; Brasil.*

Citation: Temponi EF, Lamego LLS, Costa EL, Passos GA, Baumfeld TS, Marinho LS, Campos TVO, Silva RB, Miranda VR. Bacterial colonization in orthopedic surgical tourniquets: a multicenter study. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 4. Available from URL: http://www.scielo.br/aob.

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Hospital Madre Teresa, Belo Horizonte, MG, Brazil.
Correspondence: Eduardo Frois Temponi. Hospital Madre Teresa, Avenida Raja Gabáglia, 1002 Av. Raja Gabáglia, 1093 - Luxemburgo - Gutierrez, 30441-070, Belo Horizonte, Minas Gerais, Brasil dufrois@gmail.com

Article received on 12/08/2025 approved on 01/23/2026.
Handling Editor: Camilo Partezani Helito

<< SUMÁRIO

INTRODUCTION

The use of tourniquets in orthopedic surgeries is a widely established practice recognized for its role in reducing intraoperative bleeding and improving the visibility of the surgical field[1-3]. However, reusable devices, with often insufficient cleaning between procedures, have been associated with high rates of microbial colonization[1,3-5]. Even without direct contact with the operative field, tourniquets maintain close contact with the skin and adjacent tissues, which may allow the transmission of pathogenic microorganisms.

International studies report contamination rates between 68% and 96% in reusable tourniquets, including the presence of *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and multidrug-resistant microorganisms[2,3,5-7]. In contrast, sterile single-use devices exhibit colonization rates close to zero[3,8,9]. Ahmed et al.[4] demonstrated that disinfection with chlorhexidine wipes can reduce the microbial load by up to 99% in tourniquets. Szymczyk et al.[5], in turn, identified averages of up to 545 CFU/cm² in reused devices, especially in emergency services.

Despite the importance of the topic, there is a lack of national data documenting the magnitude of the risk in our country[9]. Moreover, there is an absence of specific regulations from agencies such as the National Health Surveillance Agency (ANVISA), which contributes to the heterogeneity in disinfection practices adopted among different hospital institutions[10].

In light of this scenario, the present multicenter study aims to evaluate the prevalence and microbiological profile of microbial load colonization in reusable tourniquets used in orthopedic surgeries in Brazil[11]. It also seeks to discuss the implications of these findings for healthcare-associated infection control, proposing evidence-based strategies that may underpin institutional policies and national biosafety guidelines.

METHODOLOGY

Design and Ethical Aspects

This was an observational, cross-sectional, and multicenter study with a quantitative approach, conducted in conducted in six Brazilian hospitals. The participating hospitals were coded for analysis and are not individually identified in the results. The study was approved by the Institutional Research Ethics Committee through the Plataforma Brasil system, under the substantiated opinion No. 7.621.206 and CAAE 89171425.0.0000.51271.

Sample Calculation

Based on Thompson et al.[3] and Ahmed et al.[4], which reported contamination rates between 68% and 96%, we assumed an expected prevalence of 80% for bacterial colonization. Considering a margin of error of 10% and a confidence interval of 95%, a minimum sample of 50 tourniquets was estimated to ensure adequate descriptive power.

Sample Collection

Fifty-four reusable tourniquets used in orthopedic surgeries (elective and emergency) between May and July 2025 were analyzed. The collection followed consecutive sampling and was conducted immediately after surgical use and before the routine institutional disinfection procedure.

A sterile swab was used for each device, rubbed in a rotational and unidirectional manner over the internal surface of the tourniquet (the area in contact with the skin). The collection protocol standardized the friction over a linear extent of 10 cm. Considering the average width of the swab tip of approximately 1 cm, the total sampled area was estimated at 10 cm² for the purposes of calculating bacterial density and benchmarking (CFU/cm²). To ensure standardization and reliability, all collectors were previously trained according to a detailed operational protocol, which included specific instructions regarding the area to be rubbed, the contact time, and the aseptic handling of the material, minimizing inter-observer variations and potential biases related to the collection technique. The time between collection and laboratory processing was standardized, limited to a maximum of two hours, in order to ensure the microbiological viability of the samples.

Microbiological Processing

The samples were inoculated onto 5% blood agar plates and MacConkey agar plates (Oxoid™), incubated at 37°C for 24 to 48 hours under aerobic atmosphere. The identification of microorganisms was performed using automated methods (VITEK® 2 Compact, bioMérieux or MALDI-TOF, Bruker®). The quantification of bacterial load was conducted through direct counting of Colony Forming Units (CFU). The microbiologists responsible for reading were blinded to the type of hospital from which the samples originated.

Data Analysis

The data were analyzed using R software (version 4.3.1). The normality of the quantitative variables was assessed using the Shapiro-Wilk test. Due to the non-parametric distribution of the data, the bacterial load (CFU) was described by median and interquartile range (IQR: Q1-Q3). The categorical variables were described in absolute and relative frequencies.

To compare proportions between public and private hospitals, Fisher's exact test was used. To compare bacterial loads (non-parametric continuous variables), the Mann-Whitney test was employed. A significance level of 5% ($p<0.05$) was adopted for all analyses.

RESULTS

Out of the 54 analyzed tourniquets, 38 exhibited microbial growth, corresponding to a prevalence of 70.4% (95% CI: 57.4% – 81.1%). There was no microbial growth in 16 tourniquets (29.6%).

Microbiological Profile

The frequency of isolated microorganisms is detailed in Table 1 (Frequency of isolated microorganisms). The microorganism most frequently isolated was coagulase-negative Staphylococcus, present in 26 samples (48.1%). Pathogens of clinical relevance were identified, including *Staphylococcus aureus* (18.5%), *Bacillus sp.* (14.8%) and *Pseudomonas aeruginosa* (9.2%). In addition to bacteria, yeasts of the genus *Candida spp.* were isolated in 2 samples (3.7%). In two tourniquets, co-colonization by more than one microorganism was detected simultaneously (Table 1).

Bacterial Load

Because bacterial load was not normally distributed of microbial load (Shapiro-Wilk $p<0.05$). The estimated overall median was 101 CFU per device (Interquartile Range [IQR]: 0 – 153). When adjusted for the sampled surface area (10 cm²), the median density was 10.1 CFU/cm².

Comparison between Institutions

Of the 28 tourniquets from public hospitals, 22 exhibited contamination (78.6%). Among the 26 tourniquets from private hospitals, 16 were contaminated (61.5%). Despite the absolute percentage

Table 1. Frequency of isolated microorganisms.

Isolated Microorganism	n (isolates)	Frequency (%)*
Coagulase-negative Staphylococcus	26	48.1%
Staphylococcus aureus	10	18.5%
Bacillus sp.	8	14.8%
Pseudomonas aeruginosa	5	9.2%
Candida spp.	2	3.7%

difference, this variation did not reach statistical significance (p=0.081; Fisher's exact test). (Figures 1 and 2)
Regarding microbial load, devices from public hospitals had a median of 117 CFU (IQR: 85 – 184), while those from private hospitals had a median of 73 CFU (IQR: 0 – 148). The comparison of bacterial load distributions between the two groups also did not demonstrate a statistically significant difference (p=0.412; Mann-Whitney test).

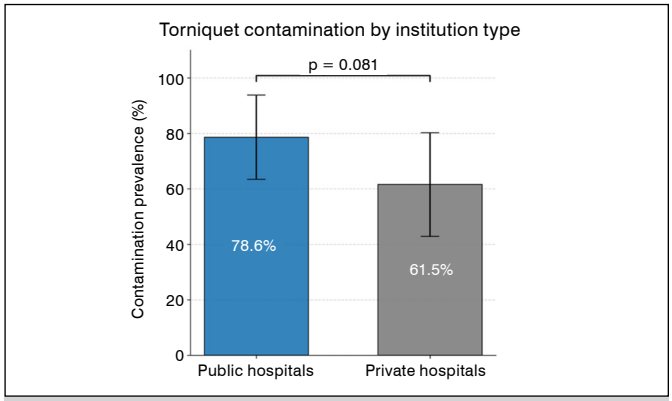


Figure 1. Prevalence of contamination by type of hospital.

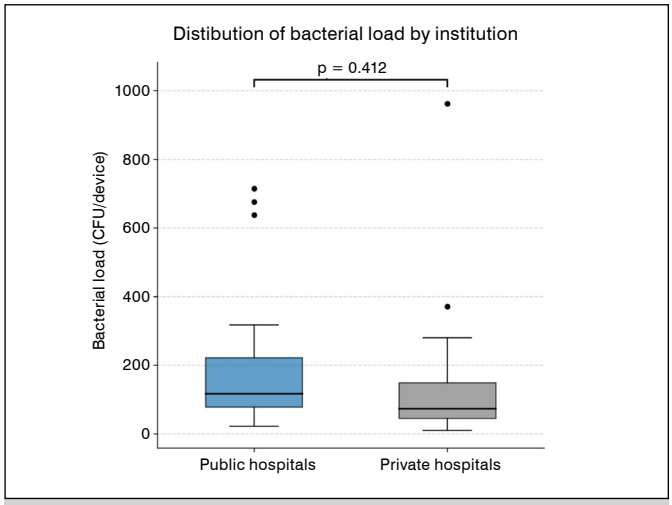


Figure 2. Distribution of bacterial load (CFU) by type of hospital.

DISCUSSION

The results of this multicenter study reveal a prevalence of microbial colonization of 70.4% in reusable orthopedic tourniquets, corroborating international data that report contamination rates ranging from 68% to 96% in similar devices[1,3-5]. The detection of a global median of 101 CFU per device (estimated at ~10.1 CFU/cm²), immediately after intraoperative use and before reprocessing, evidences a substantial biological load that challenges the current disinfection protocols. In contrast, studies demonstrate that sterile single-use devices or those subjected to sterilization protocols exhibit colonization rates close to zero, highlighting the role of reusable tourniquets as frequently overlooked high-touch contact surfaces[3]. The identified microbiological profile exceeds the harmless commensal microbiota, with a high prevalence of *coagulase-negative Staphylococcus* (48.1%) and *Staphylococcus aureus* (18.5%), aligning with recent findings by Szymczyk et al.[5]. This scenario represents a high clinical risk, as these agents are the main causes of periprosthetic infections, and the literature warns of the potential for these devices to act as reservoirs for multidrug-resistant organisms, perpetuating cycles of hospital infection[6]. Additionally, the presence of *Pseudomonas aeruginosa* (9.2%) and *Bacillus* sp. (14.8%) suggests failures in the biosafety barrier and environmental persistence of biofilm-forming organisms, which are difficult to eradicate through simple manual cleaning, while the isolation of *Candida spp.* (3.7%) reinforces the complexity of contamination favored by moisture retained in the tissues of the cuffs.
The association between the use of tourniquets and the increased risk of Surgical Site Infection (SSI) is well documented, with meta-analyses indicating a high risk in total knee arthroplasties[1,8,12]. Although factors such as tissue hypoxia have historically been blamed, recent evidence demonstrates that the tourniquet also affects the local tissue concentration of prophylactic antibiotics, compromising the effectiveness of perioperative prophylaxis[2]. In addition, the physical presence of a source rich in viable pathogens in the immediate vicinity of the surgical field constitutes a critical modifiable risk vector. We recognize as a limitation of this cross-sectional study the absence of "pre-use" immediate collection, which prevents the unequivocal distinction of whether the recovered microbial load originates from the skin flora of the current patient or represents residual contamination from previous procedures. However, from the perspective of biosafety, this distinction becomes secondary in light of the risk of cross-contamination for the next patient. The documented bacterial load (median of 117 CFU in public hospitals and 73 CFU in private ones) represents the biological challenge that the subsequent cleaning process must eliminate. Ahmed et al.[4] demonstrated that, although cleaning with chlorhexidine may reduce the load by 99%, the effectiveness in routine practice is inconsistent and operator-dependent. If reprocessing is ineffective, the identified pathogens will persist, transforming the tourniquet into a passive vector for the subsequent surgical case. The comparison between public and private hospitals did not demonstrate a statistically significant difference in contamination rates (p=0.081) or in the median bacterial load (p=0.412), clinically suggesting that the contamination of tourniquets is a systemic problem in orthopedic practice, regardless of management model or resource availability. The persistence of contamination above 60% in both groups indicates that current protocols based on manual cleaning are insufficient. The adoption of alternative materials, such as silicone tourniquets, has demonstrated a reduction in contamination rates compared to traditional fabric ones, presenting itself as a viable alternative to reduce microorganism adhesion[10]. Furthermore, systematic reviews reinforce that the concern for the safe use of the tourniquet should be universal, encompassing everything from pediatrics to complex reconstructive surgeries[7]. Although the acquisition of disposable sterile tourniquets represents a higher initial cost, the economic analysis should consider the burden of treating complications, as the clinical impact of failures in arthroplasties justifies preventive investments[13]. The aggregated cost of treating a single deep infection exponentially exceeds the investment in adopting single-use devices, in addition to mitigating the spread of resistant microorganisms in the hospital environment[9,11]. For institutions where the transition to disposables is not immediate, it is recommended to validate high-level disinfection protocols, the mandatory use of waterproof sterile protection (*stockinette*) under the tourniquet, and periodic monitoring of microbial load.
Despite the relevance of the findings for national biosafety, this study presents limitations that must be considered in the interpretation of the results. The main restriction lies in the cross-sectional design without an immediate "pre-use" baseline collection, which prevents the unequivocal distinction between contamination arising from the skin microbiota of the current patient and the residual load resulting from failures in the reprocessing of previous surgeries. Although this limitation restricts the precise definition of the origin of the inoculum,

it does not invalidate the finding of the risk of cross-contamination, as the detected biological load represents the real challenge to be eliminated before the next use. Additionally, the swab sampling technique, although standardized, may underestimate the total microbial load, especially those deeply adhered to the fabric of the cuff or organized in biofilms, which would be better recovered by sonication methods, logistically unfeasible in the proposed multicenter design.

Another important limitation was the lack of antimicrobial sensitivity testing (antibiogram), which hindered the characterization of the resistance profile of the isolates, such as the prevalence of *S. aureus* resistant to methicillin (MRSA). The absence of negative field controls (swabs exposed to the environment without contact with the device) also prevents the exact quantification of background environmental contamination, although the rigor in the aseptic technique aimed to minimize this bias. Moreover, clinical covariate variables, such as the duration of surgery, type of antiseptic skin preparation, or the use of protective meshes (*stockinettes*) under the tourniquet, factors that could influence the final colonization density, were not collected. Finally, as the study focused on the colonization of the device and did not follow patients longitudinally, it is not possible to establish a direct causal correlation between the contamination of the tourniquets and the rates of surgical site infection (SSI) in the participating institutions, leaving this inference based on biological plausibility and comparative literature.

CONCLUSION

This multicenter study highlights a high prevalence of microbial colonization (70.4%) in reusable orthopedic tourniquets in Brazil, with the identification of microorganisms of high clinical relevance, including *Staphylococcus aureus*, *Pseudomonas aeruginosa*, and *Candida* spp. The absence of a statistically significant difference in contamination rates and microbial load between public and private hospitals indicates that the insufficiency of reprocessing protocols is a systemic and cross-sectional challenge in national orthopedic practice, not limited to scenarios with resource constraints. It is concluded that reusable tourniquets, when subjected only to conventional cleaning, act as potential reservoirs of pathogens in the surgical environment. In light of these findings, it is recommended to review institutional biosafety guidelines, prioritizing the adoption of sterile disposable tourniquets, especially in implant surgeries, or, in the impossibility of this transition, the rigorous implementation of high-level disinfection associated with the mandatory use of waterproof sterile barriers.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. EFT: writing, conceptualization, supervision; LLSL: data analysis, methodology,
investigation; ELC: article review and intellectual contribution; GAP: supervision, writing, and resources; TSB: validation, review, writing; LSM: visualization, investigation,
writing; TVOC: conceptualization, data analysis; RBS: methodology, investigation; VRM: programming, methodology, investigation, analysis, and resources.

DATA AVAILABILITY DECLARATION

The underlying contents of the research text are contained in the manuscript.

REFERENCES

1. Magan AA, Dunseath O, Armonis P, Fontalis A, Kayani B, Haddad FS. Tourniquet use in total knee arthroplasty and the risk of infection: a meta-analysis. J Exp Orthop. 2022;9(1):62. doi: 10.1186/s40634-022-00485-9.
2. Montreuil J, Tanzer M, Zhang YL, Rajda E, Avizonis D, Hart A. Tourniquet use and local tissue concentrations of antibiotics and risk of surgical site infections. JAMA Netw Open. 2024;7(8):e2429702. doi: 10.1001/jamanetworkopen.2024.29702.
3. Thompson SM, Middleton M, Farook M, Cameron-Smith A, Bone S, Hassan A. The effect of sterile versus non-sterile tourniquets on microbiological colonisation in lower limb surgery. Ann R Coll Surg Engl. 2011;93(8):589-90. doi: 10.1308/147870811X13137608455334.
4. Ahmed SMY, Ahmad R, Case R, Spencer RF. A study of microbial colonisation of orthopaedic tourniquets and the effectiveness of cleaning methods. Ann R Coll Surg Engl. 2009;91(2):131-4. doi: 10.1308/003588409X359402.
5. Szymczyk J, Kurpas M, Krasiński B, Zorena K, Mędrzycka-Dąbrowska W. Reusable tourniquets as potential transmitters of infection: a microbiological analysis. Microorganisms. 2025;13(1):152. doi: 10.3390/microorganisms13010152.
6. Szymczyk J, Månsson M, Mędrzycka-Dąbrowska W. Reusable tourniquets for blood sampling as a source of multi-resistant organisms: a systematic review. Front Public Health. 2023;11:1258692. doi: 10.3389/fpubh.2023.1258692.
7. Pintar V, Brookes C, Trompeter A, Bridgens A, Hing C, Gelfer Y. A systematic review of tourniquet use in paediatric orthopaedic surgery: can we extrapolate from adult guidelines? EFORT Open Rev. 2024 Jan 9;9(1):80-91. doi: 10.1530/EOR-23-0091. PMID: 38193578; PMCID: PMC10823570.
8. Sun C, Zhang X, Ma Q, Tu Y, Cai X, Zhou Y. Impact of tourniquet during total knee arthroplasty when tranexamic acid was used: a meta-analysis of randomized controlled trials. J Orthop Surg Res. 2022;17(1):18. doi: 10.1186/s13018-021-02898-1.
9. Natarajan A, Das S, Chaudhary N. Microbial profile of reusable phlebotomy tourniquets at a rural tertiary-care teaching hospital. Cureus. 2023;15(11):e49328. doi: 10.7759/cureus.49328.
10. Grohmann M, Schomakers L, Wolschendorf F, Grosch J, Lindner S, Witte AK. Reduced bacterial contamination rates detected on silicone tourniquets compared to conventional tourniquets in clinical routine. BMC Infect Dis. 2020;20(1):247. doi: 10.1186/s12879-020-04975-y.
11. Leitch A, McCormick I, Gunn I, Gillespie T. Reducing the potential for phlebotomy tourniquets to act as a reservoir for methicillin-resistant Staphylococcus aureus. J Hosp Infect. 2006;63(4):428-31. doi: 10.1016/j.jhin.2006.03.006.
12. Alcelik I, Pollock RD, Sukeik M, Bettany-Saltikov J, Armstrong PM, Fismer P. A comparison of outcomes with and without a tourniquet in total knee arthroplasty: a systematic review and meta-analysis of randomized controlled trials. J Arthroplasty. 2012;27(3):331-40. doi: 10.1016/j.arth.2011.03.040.
13. Yi Z, Yan L, Haibo S, Yuangang W, Mingyang L, Yuan L, et al. Effects of tourniquet use on clinical outcomes and cement penetration in TKA: randomized controlled trial. BMC Musculoskelet Disord. 2021;22(1):126. doi: 10.1186/s12891-021-03968-5.

DOI: http://dx.doi.org/10.1590/1413-785220263404e300122

RADIOFREQUENCY THERAPY IN PATIENTS WITH OSTEOARTHRITIS AWAITING KNEE ARTHROPLASTY

TERAPIA DE RADIOFREQUÊNCIA EM PACIENTES COM OSTEOARTRITE AGUARDANDO ARTROPLASTIA DE JOELHO

GLAUCUS CAJATY DOS SANTOS MARTINS[1], *INGRID CARVALHO ANDRADE*[1], *ALEXANDRE PEIXOTO DE MELLO*[2], *BRUNO PARILHA COUTINHO*[1], *VINICIUS SCHOTT GAMEIRO*[3]

1. Hospital Federal de Ipanema, Servico de Ortopedia e Traumatologia, Rio de Janeiro, RJ, Brazil.
2. Universidade Federal do Estado do Rio de Janeiro (Unirio), Faculdade de Medicina, Rio de Janeiro, RJ, Brazil.
3. Universidade Federal Fluminense (UFF), Faculdade de Medicina, Niterói, RJ, Brazil.

ABSTRACT

Objective: To evaluate the palliative use of conventional radiofrequency (RF) in patients awaiting total knee arthroplasty (TKA). Methods: Seventy-one knees with osteoarthritis and indication for TKA were evaluated and divided into two groups: Group 1 with 51 knees treated with RF and Group 2 with 20 patients treated with intra-articular corticosteroids. The primary outcome was pain response using the visual analog scale (VAS) at six months postoperatively. The secondary outcome was functional assessment using the WOMAC and KSS scales. The MCID (minimal clinically important difference) criterion was used. The influence of clinical and radiographic parameters on the procedural outcome was evaluated. Results: The RF group showed improvement in VAS at 6 months compared to the control group ($\chi 2$ test, p=0.0124). The RF group obtained a superior result than the control group using the WOMAC scale ($\chi 2$ test, p=0.044). The KSS, divided into knee and function subgroups, performed better than the control group (p=0.02 and p=0.048). Knee angulation and degree of osteoarthritis were associated with procedural prognosis ($\chi 2$ test, p=0.028, p=0.019). Conclusion: The RF group showed superior pain and functional improvements compared to the control group. Osteoarthritis severity and knee angulation influenced functional and pain outcomes. ***Level of Evidence II; Comparative Prospective Study.***

Keywords: Corticoid; Knee; Osteoarthritis; Radiofrequency.

RESUMO

Objetivo: Avaliar o uso paliativo de Radiofrequência (RF) convencional em pacientes que aguardam o procedimento de artroplastia total de joelho (ATJ). Métodos: Avaliados 71 joelhos com osteoartrite com indicação de ATJ divididos em 2 grupos: Grupo 1 com 51 joelhos tratados com RF e Grupo 2 com 20 pacientes tratados com corticoide intra-articular. O desfecho primário foi a resposta da dor pela escala visual analógica (EVA) ao sexto mês pós-operatório. O desfecho secundário foi a avaliação funcional pelas escalas WOMAC e KSS. Utilizado o critério de MCID (minimal clinically important difference). Avaliou-se a influência de parâmetros clínicos e radiográficos no resultado do procedimento. Resultados: O grupo RF comparado ao controle teve melhora do EVA ao final de 6 meses, teste $\chi 2$ (p=0,0124). O grupo RF obteve resultado superior ao controle pela escala WOMAC (teste $\chi 2$, p=0,044). O KSS dividido em subgrupos joelho e função obteve desempenho superior ao controle (p=0,02 e p=0,048). Angulação do joelho e grau de osteoartrite mostraram relação com prognóstico do procedimento (teste $\chi 2$, p=0,028, p=0,019). Conclusão: O grupo RF apresentou melhora da dor e funcional superiores ao grupo controle. Gravidade da artrose e angulação do joelho influenciaram os resultados funcionais e da dor. ***Nível de Evidência II; Estudo Prospectivo Comparativo.***

Descritores: *Corticoide; Joelho; Osteoartrite; Radiofrequência.*

Citation: Martins GCS, Andrade IC, Mello AP, Coutinho BP, Gameiro VS. Radiofrequency Therapy in Patients with Osteoarthritis Awaiting Knee Arthroplasty. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 6. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

The demand for knee arthroplasty procedures (TKA) in the public health system has progressively increased over the last few decades[1-4]. Various therapeutic modalities, including the use of radiofrequency therapy (RF) in the knee[3-10], have been utilized palliatively while awaiting TKA.

Radiofrequency therapy in the knee involves the use of thermal energy generated by an electrode acting at the level of the genicular nerves. This leads to functional deactivation, preventing the transmission of painful stimuli from joint areas of the knee from reaching the cerebral sensory cortex, thereby blocking the sensation of pain [6-9].

All authors declare no potential conflict of interest related to this article.

The study was conducted at Hospital Federal de Ipanema, Orthopedics and Traumatology Service, Rio de Janeiro, RJ, Brazil.
Correspondence: Glaucus Cajaty dos Santos Martins. 29, Rua Antônio Parreiras, 67-69,Ipanema, Rio de Janeiro, RJ, Brazil. CEP 22411-020.

Article received on 08/23/2025 approved on 03/10/2026.
Handling Editor: Camilo Partezani Helito

In meta-analysis studies[11-14], RF has proven effective in pain control for knee osteoarthritis in patients with poor clinical conditions to undergo TKA[1-9]. In public health, the waiting time for TKA can reach up to three years. This period inflicts considerable suffering on the patient. Palliative pain measures need to be adopted until the definitive surgery[4,5].

In the reviewed literature, no studies were identified that exclusively address the use of RF in individuals indicated for TKA but who do not undergo it due to unavailability of the prosthesis.

A comparison was made between RF and the classical treatment of intra-articular corticosteroid injection[10] in individuals with OA indicated for TKA. Parameters of the Visual Analog Scale for Pain (VAS) and functional WOMAC and KSS were evaluated[15,16].

It should be emphasized the original aspect of this work in our environment and the potential contribution of the research results to the adoption of this palliative therapy in patients awaiting TKA.

METHODOLOGY

This research adhered to the guidelines of the Helsinki Declaration and was approved by the research ethics committee of our hospital (CEP 5.071.669). This article constitutes a comparative and prospective cohort study.

A total of 61 individuals (55 women and 6 men) were selected from the knee outpatient clinic of a public institution with an indication for TKA. There was an expectation to wait at least one year on the waiting list.

The RF group was formed by 51 knees, and another 20 participants constituted the control group with intra-articular corticosteroid treatment. The research was conducted from January 2022 to December 2023. The authors initially sought to select the study and control groups in a 2:1 ratio.

Inclusion criteria: individuals over 50 years of age, with radiographic primary osteoarthritis, a visual analog scale (VAS) score of 5 or higher, no improvement with conservative treatment, and no previous open surgeries on the knee.

Exclusion criteria: rheumatological diseases, radiculopathies, use of anticoagulants, previous intra-articular injections with corticosteroids or hyaluronic acid within a period of less than six months.

The sample was based on works by El-Hakeim et al[3] and Konya et al[17] which determined that 30 patients would be sufficient to detect changes in the VAS (primary outcome) in 80% of cases with a significance level of 5%.

Pain was measured using the visual analog scale (VAS) from 0 to 10. Success was considered as a result of a reduction in pain equal to or greater than 50% after treatment.

Functional impairment was measured using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[15] which evaluates three domains: pain; stiffness; function. The minimal clinically important difference (MCID) criterion was employed (*minimal clinically important difference*)[15], which indicates the minimum value representing a substantial improvement perceived by the patient. In the case of WOMAC and its subgroups, this would represent a variation of 16% in relation to the value obtained prior to the procedure. In the case of a successful outcome, the functional improvement value of the MCID must be achieved or exceeded.

The Knee Society Score (KSS)[16] was also employed. This scale combines subjective and objective information and separates the knee score (which evaluates pain, stability, and range of motion) from the functional score (ability to walk, ascend, and descend stairs). The MCID was the minimum improvement of 9 points in the knee criterion and 10 points in the function criterion[16].

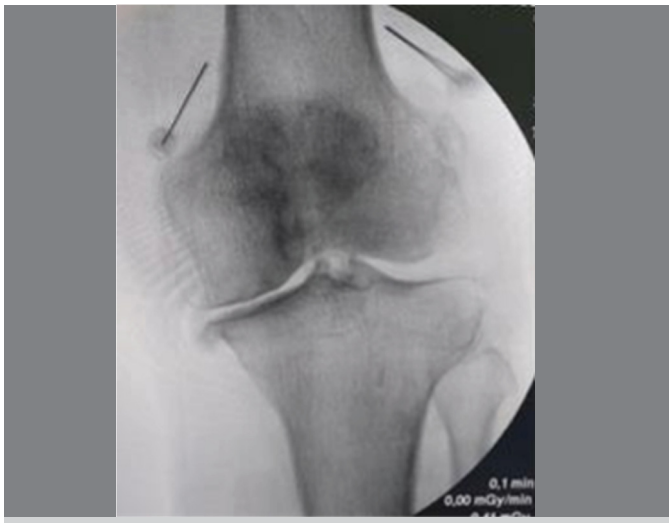
Pain according to the VAS scale was assessed at the pre-intervention moment, the first month, and the sixth month after the procedure.

A comparison was made between the control and RF groups. The WOMAC and KSS scales were evaluated similarly to the VAS. Data were collected regarding age, gender, body mass index (BMI), duration of the disease, radiographic assessment of osteoarthritis according to the Kelgren and Alhback classifications, tibiofemoral radiographic alignment, and pre-treatment pain intensity. The same parameters were analyzed in the RF group regarding their prognostic value concerning the results. To allow for categorical statistical evaluation, types I to III of the Kelgren classification were grouped as mild or moderate involvement, and type IV as severe. Based on a previous study[18], types I to III of Ahlback and types IV and V were divided into two different groups, which would present distinct morphopathological changes. A tibiofemoral alignment of 10 degrees or more was considered pathological valgus, neutral between 1 and 9 degrees, and varus at or below zero degrees. Due to there being only three cases in the neutral category, these were discarded for better statistical significance. The analysis and collection of results were performed by a physician not involved in the procedures.

- Corticosteroid Injection Technique: following classical methodology in the literature with the use of depot corticosteroid[10].
- RF Technique: With methodology already described in the literature[1-5], conventional cannulas were used. The patient was in a supine position, asepsis and antisepsis were performed, operative fields were placed, the knee was flexed at 15 degrees, under image intensifier control, infiltration of lidocaine was performed in the skin, subcutaneous tissue, muscle, and periosteum of the medial and lateral metaphyseal-diaphyseal junction of the distal femur and medial proximal tibia (Figures 1, 2, and 3). The inferior lateral genicular nerve is not approached due to the risk of compromising the fibular nerve. RF cannulas of 10 cm x 22 Gauge were used. Sensory stimulation test at 50Hz with a threshold below 0.6V. To avoid motor impairment, stimulation was performed at 2Hz and 2.0V, and muscle contractions should not occur. Through the cannulas, 2ml of lidocaine was infiltrated, and the cannula was heated to 70 degrees Celsius for 150 seconds.

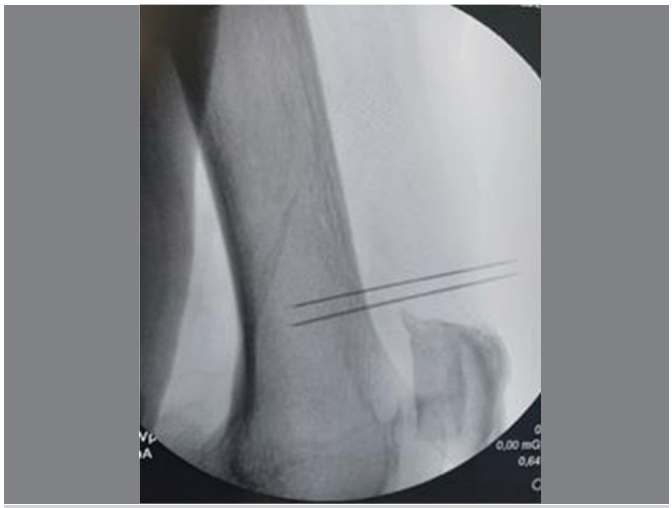
Statistical Analysis

The comparison of baseline variables between the groups (RF and control) was analyzed using the Student's t-test for independent samples, Mann-Whitney for numerical variables, and the $\chi 2$ test



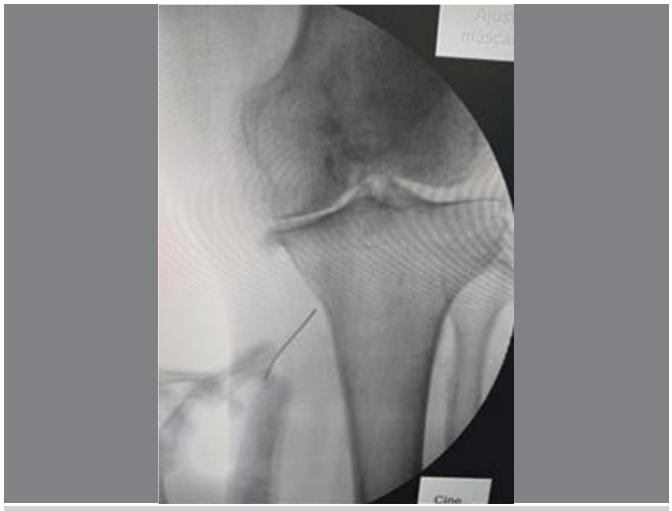
Source: Author's collection (2024).

Figure 1. Positioning in AP radiography for thermoregulation of the superior medial genicular nerve and superior lateral genicular nerve. AP x-ray incidence of medial and lateral superior genicular nerves radiofrequency.



Source: Author's collection (2024).

Figure 2. Positioning in Profile radiography for thermoregulation of the superior medial genicular nerve and superior lateral genicular nerve. rofile x-ray incidence of medial and lateral superior genicular nerves radiofrequency.



Source: Author's collection (2024).

Figure 2. Positioning in AP incidence for thermoregulation of the inferior medial genicular nerve. AP x-ray incidence of medial inferior genicular nerve radiofrequency.

or Fisher's exact test for categorical variables. The comparison of treatment response (success and failure) between the groups was analyzed using the $\chi 2$ test.

RESULTS

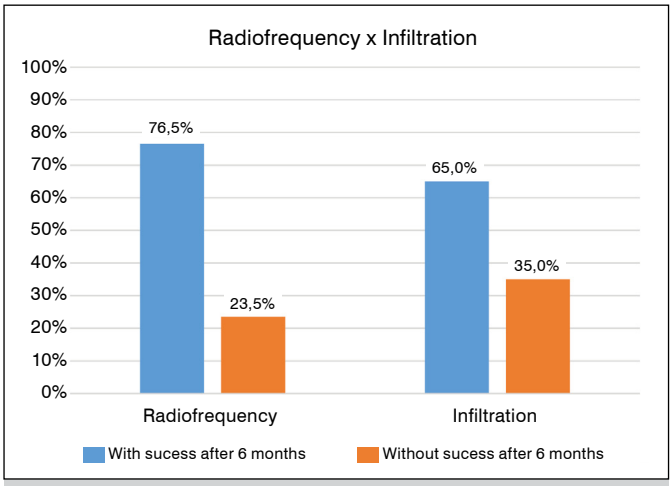
The RF group exhibited a higher prevalence of individuals with a tendency towards valgus angle than the control group (17.6% vs 0%), p = 0.033. There was no statistical difference in the other variables between the groups. The RF group showed success on the VAS scale in the 1st month (74.5%) and in the 6th month (76.5%), significantly higher than the infiltration group (30.0% and 35.0%, p=0.0005 and p=0.0001) respectively. In the RF group, the average VAS varied from 8.8 to 3.2 (63% variation), while the IAC group varied from 8.9 to 5.8 (35%) at the 6th month. The RF group was superior to the control group with an odds ratio of 2.18 (Table 1 and Figure 4). The RF group demonstrated success on the WOMAC at the 1st month (90.2%) and at the 6th month (90.2%) greater than the infiltration group (70.0%, 70.0%), with p=0.044, p=0.044.

Table 1. Success of the treatment in the complete sample (VAS and WOMAC) and in the subgroups of WOMAC.

Variable	Total		Radiofrequency		Infiltration		*p value*	RR
	n	%	n	%	n	%		
VAS Success 0–1 month								
yes	44	62.0	38	**74.5**	6	30,0	**0.0005**	**2.48**
no	27	38,0	13	25,5	14	70,0		
VAS Success 0–6 months								
yes	46	64,8	39	**76.5**	7	35.0	**< 0.0001**	**2.18**
no	25	35.2	12	23.5	13	65.0		
WOMAC Success 0–1 month								
yes	60	84,5	46	**90.2**	14	70.0	**0.044**	**1.29**
no	11	15,5	5	9.8	6	30,0		
WOMAC Success 0–6 months								
yes	60	84,5	46	**90.2**	14	70.0	**0.044**	**1.29**
no	11	15.5	5	9.8	6	30.0		
WOMAC Pain Success 0–6 months								
yes	61	85,9	47	**92.2**	14	70.0	**0.024**	**1.32**
no	10	14.1	4	7.8	6	30.0		
WOMAC Stiffness Success 0–6 months								
yes	13	18,3	13	**25.5**	0	0.0	**0.008**	*Undefined*
no	58	81,7	38	74.5	20	100.0		
WOMAC Function Success 0–6 months								
yes	58	81,7	44	86.3	14	70.0	0.10	1.23
no	13	18.3	7	13.7	6	30.0		

Data are expressed as frequency (n) and percentage (%). Chi-square test or Fisher's exact test. RR: relative risk of radiofrequency for success.

Source: Own authorship (2024).



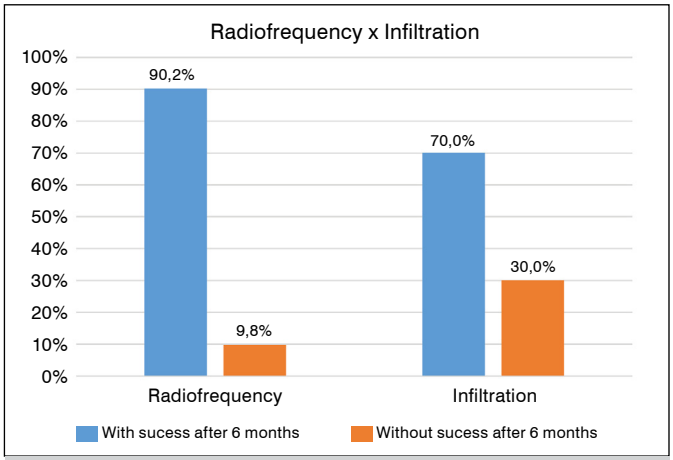
Source: Author's collection (2024).

Figure 4. Comparative result of the VAS between the study and control groups after 6 months.

In the WOMAC subgroups, pain at the 6th month (92.2%), WOMAC stiffness at the 6th month (25%) showed superior performance compared to the control group (70.0% and 0%, p=0.024 and p=0.008) respectively (Figure 5). The WOMAC in the RF group decreased from 63.3 to 42.5 (37.5% percentage variation) in 6 months, while group 2 varied from 60.5 to 46.3 (20.96%). The RF group exhibited a success rate according to the KSS Knee at the 6th month (98.0%) and KSS Function at the 6th month (78.4%) significantly higher than the infiltration group (80.0% and 55.0%, p=0.020, and p=0.048) respectively (Table 2, Figure 6).

In the RF group, a high BMI (>30) correlated with a favorable prognosis with improvement on the VAS scale and the WOMAC scale subgroup regarding pain response at the end of 6 months (X2 test p=0.03 and p=0.019, respectively).

Knee angulation in valgus demonstrated a negative response to the RF method when evaluated by the VAS scale, total WOMAC, and WOMAC in the pain subgroup (X2 test p=0.032, p=0.040 and p=0.028, respectively). The degree of osteoarthritis assessed by the Kelgren and Ahlback scales showed a relationship between increased radiographic severity and improvement in the stiffness



Source: Author's collection (2024).

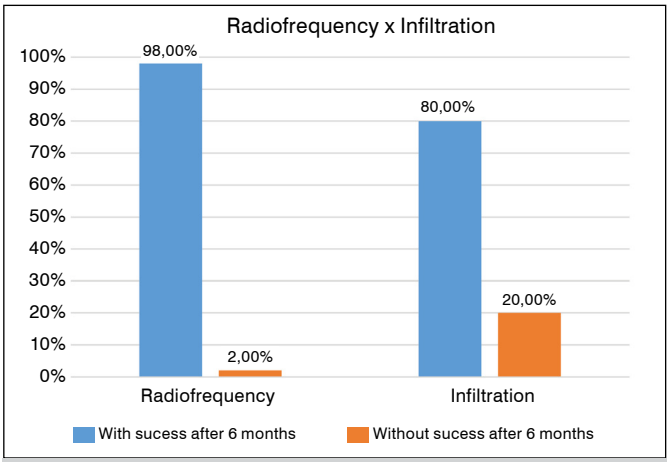
Figure 5. Comparative result of the WOMAC between the study and control groups after 6 months.

Table 1. Treatment success in the complete series relative to KSS Knee and KSS Function.

Variable	Total		Radiofrequency		Infiltration		*p value*	RR
	n	%	n	%	n	%		
KSS Knee Success 0–1 month								
yes	66	93.0	50	**98.0**	16	**80.0**	**0.020**	**1.22**
no	5	7.0	1	2.0	4	20.0		
KSS Knee Success 0–6 months								
yes	66	93.0	50	**98.0**	16	**80.0**	**0.020**	**1.22**
no	5	7.0	1	2.0	4	20.0		
KSS Function Success 0–1 month								
yes	49	69.0	38	74.5	11	55.0	0.11	1.35
no	22	31.0	13	25.5	9	45.0		
KSS Function Success 0–6 months								
yes	51	71.8	40	**78.4**	11	**55.0**	**0.048**	**1.42**
no	20	28.2	11	21.6	9	45.0		

Data are expressed as frequency (n) and percentage (%). Chi-square test or Fisher's exact test. RR: relative risk of radiofrequency for success.

Source: Author's arquives



Source: Author's collection (2024).

Figure 5. Comparative result of the KSS Knee between the study and control groups after 6 months.

component response of the WOMAC at the 6th month (X2 test p=0.019, p=0.037, respectively). The other parameters did not have prognostic relevance. The remaining parameters did not show prognostic relevance.

It was noted that in the RF group, non-obese individuals had a significantly higher prevalence of genu valgum cases (p<0.001) than obese individuals. As previously shown, this deformity would be linked to inferior functional outcomes. The good results of the obese group could be explained by the lower prevalence of genu valgum cases and not necessarily by the BMI itself.

No clinical complications were detected with the RF technique.

DISCUSSION

The main finding of this study was that the RF group achieved better results in pain control and in the functional scales WOMAC and KSS than the group treated with corticosteroids at the sixth month of follow-up. It was evidenced that factors such as the severity of osteoarthritis and the angulation of the knee in valgus influenced the outcome of pain response and/or functionality.

In this study, the main indication for RF occurred in patients selected for TKA, who were fit and willing to undergo surgery, however, without the availability of the implant. Such a situation creates a bias that challenges the RF method, as the patient is fully aware that there is a gold standard, namely, the prosthesis.

No test anesthetic block was used to determine the indication for the RF procedure due to its low prognostic value[19,20].

The authors used the image intensifier to perform the procedures. There was a disadvantage of exposure to radioactivity compared to the use of ultrasound. Kim et al[14] did not demonstrate differences in outcomes between the techniques.

Juni et al[10] evidenced improvement in pain up to the sixth week after corticosteroid infiltration for knee arthritis. After six months, there would be no influence on the maintenance of joint space or improvement in quality of life.

In 2016, Sari et al[1] conducted the first randomized study on knee arthritis using conventional RF compared to the infusion group with corticosteroids. At the end of three months, the RF group showed improvement in pain (VAS) and a decrease in WOMAC.

Davis et al[2] in 2018 compared cooled RF with intra-articular corticosteroid infiltration in knee arthritis. There were 76 treated with RF and 75 with corticosteroids. At the sixth month, a reduction of more than 50% in pain intensity occurred in 74.1% of the study group versus 16.2% of the control. In the current research, the results obtained were similar to those of Davis et al[2] and Sari et al[1]. The functional gains obtained tended to be maintained from the first to the sixth month, similar to other authors[17,21].

Recent meta-analyses[2,13] identified sustained improvement in pain levels and the WOMAC scale of the RF groups compared to the control groups (hyaluronic acid, corticosteroid, placebo), not reporting serious adverse effects similar to the present study. In this research, the authors utilized conventional cannulas. The cooled RF tip provides a volumetric lesion 5 to 20 times greater than the traditional tip[22,23]. Some authors[22,24] argue that cooled RF would achieve better and more lasting results than the conventional method. However, a meta-analysis[25] addressing cooled, pulsed, and conventional RF did not demonstrate a difference in pain improvement up to 12 months post-procedure. This observation reinforces the use of conventional RF, a technique more likely to be employed in public programs for palliative pain management.

The radiographic alignment of the limb had not yet been studied in the researched literature regarding its impact on the outcome of RF therapy. It was found that patients with genu valgum would exhibit a lower response regarding pain improvement (VAS) and function (WOMAC). Genu valgum generally presents as a more complex pathology than genu varum for surgical approaches. However, in the case of knee RF, an explanation for the poorer results could be the limited number of points addressed by the method on the lateral side, which would be the predominantly painful side. In this study, only one lateral point was treated. Chen et al[25] demonstrated that the more points addressed in the knee, the better the response. The approach to the recurrent branch of the fibular nerve on the lateral tibia could be an option that improves the response to RF.

The severity of genu varum deformity was not related to the clinical response to the RF method.

In the initial analysis of the results, obesity (BMI>30) appeared to be a positive prognostic factor for pain improvement (VAS scale). Such a finding would be corroborated by Chen et al[25]. On the other hand, Santana Pineda et al[26] described elevated BMI as a factor for inferior response to RF. After comparing the obese and non-obese groups in RF, a lower prevalence of genu valgum cases was identified among the obese, which initially led to a mistaken conclusion regarding obesity and good prognosis. In fact, the smallest number of cases of valgus in the present cohort of obese individuals could be a more plausible explanation for this result. According to Kapural et al[6], the BMI would not have a relationship with the outcome of RF therapy.

This study demonstrated that individuals treated with RF and with more advanced radiographic knee arthritis showed functional improvement measured by the WOMAC stiffness scale. Sari et al[1] and Caragea et al[27] had already reported similar observations. The Alhback classification, widely used by orthopedic surgeons, allows for staging with more elaborate severity types than the Kelgren classification[18]. No other article was found regarding RF using the Alhback classification. It should also be considered that the group studied here addresses severe cases, with a long evolution and with a well-established indication for TKA.

Recently, Dias et al[28] published a randomized study conducted on patients with Kelgren IV knee osteoarthritis comparing the use of pulsed RF (17 patients) versus infiltration with phenol (18 individuals), demonstrating improvements in pain and function maintained until the third month of follow-up, without superiority of one methodology over the other.

The current research has as a critical factor the non-randomization of cases; however, the statistical analysis demonstrated that the RF and control groups were comparable and homogeneous, except for genu valgum, which was shown to be more prevalent in the RF group and related to inferior functional prognosis. Despite this, the RF group showed a response significantly superior to the control group regarding pain improvement and functional outcome. Another point would be the lack of evaluation of emotional state and opioid use. It is well established in the literature[25-27] that depression and emotional disorders negatively influence outcomes. The reliable quantification of pain medication use in patients who frequently present with multiple pain sites is a difficult and unreliable task.[6]

Positive points include the fact that the study is prospective, involving patients with an indication for TKA who wish to undergo surgery. This fact generates a more demanding view regarding the RF method, which, combined with the use of measures such as the MCID, places the results under more rigorous scrutiny. Furthermore, the evaluation of parameters not previously studied, such as tibiofemoral alignment and the Ahlback classification, deserves emphasis.

The initial results have been encouraging. The research should continue aiming at the improvement of the technique, the use of new sites for RF, and the identification of the most suitable patient for the procedure through the determination of epidemiological, clinical, and radiographic factors with prognostic influence.

CONCLUSION

The RF group exhibited greater improvements in pain and function compared to the control group. The severity of osteoarthritis and knee angulation influenced the functional and pain outcomes.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. MGCS: substantial contribution to the conception of the study, execution of procedures, and manuscript writing. AIC: contribution to data collection, writing, and revision of the manuscript. MAP: contribution to data interpretation and preparation of graphs and tables. CBP: contribution to the execution of the procedures described in the study and data acquisition. GVS: contribution to manuscript revision and final approval of the version to be published.

DATA AVAILABILITY DECLARATION

The underlying contents of the research text are contained in the manuscript.

REFERENCES

1. Sari S, Aydın ON, Turan Y, Özlülerden P, Efe U, Kurt Ömürlü İ. Which one is more effective for the clinical treatment of chronic pain in knee osteoarthritis: radiofrequency neurotomy of the genicular nerves or intraarticular injection?. Int J Rheum Dis. 2018;21(10):1772-8. doi: 10.1111/1756-185X.12925.
2. Davis T, Loudermilk E, DePalma M, Hunter C, Lindley D, Patel N, et al. Prospective, Multicenter, Randomized, Crossover Clinical Trial Comparing the Safety and Effectiveness of Cooled Radiofrequency Ablation With Corticosteroid Injection in the Management of Knee Pain From Osteoarthritis. Reg Anesth Pain Med. 2018;43(1):84-91. doi: 10.1097/AAP.0000000000000690.
3. El-Hakeim EH, Elawamy A, Kamel EZ, Goma SH, Gamal RM, Ghandour AM, et al. Fluoroscopic Guided Radiofrequency of Genicular Nerves for Pain Alleviation in Chronic Knee Osteoarthritis: A Single-Blind Randomized Controlled Trial. Pain Physician. 2018;21(2):169-77.
4. Chen AF, Khalouf F, Zora K, DePalma M, Kohan L, Guirguis M, et al. Cooled Radiofrequency Ablation Compared with a Single Injection of Hyaluronic Acid for Chronic Knee Pain: A Multicenter, Randomized Clinical Trial Demonstrating Greater Efficacy and Equivalent Safety for Cooled Radiofrequency Ablation. J Bone Joint Surg Am. 2020;102(17):1501-10. doi: 10.2106/JBJS.19.00935.
5. Chen AF, Khalouf F, Zora K, DePalma M, Kohan L, Guirguis M, et al. Cooled radiofrequency ablation provides extended clinical utility in the management of knee osteoarthritis: 12-month results from a prospective, multi-center, randomized, cross-over trial comparing cooled radiofrequency ablation to a single hyaluronic acid injection. BMC Musculoskelet Disord. 2020;21(1):363. doi: 10.1186/s12891-020-03380-5.
6. Kapural L, Lee N, Neal K, Burchell M. Long-Term Retrospective Assessment of Clinical Efficacy of Radiofrequency Ablation of the Knee Using a Cooled Radiofrequency System. Pain Physician. 2019;22(5):489-94.
7. Ariel de Lima D, Gonçalves MCK, Grando STCES, Cintra TLL, Pinto DM, Gonçalves RK. Indications of the Neurotomy of Genicular Nerves by Radiofrequency for the Treatment of Knee Osteoarthritis: A Literature Review. Rev Bras Ortop. 2019;54(3):233-40. doi: 10.1055/s-0039-1692121.
8. Davis T, Loudermilk E, DePalma M, Hunter C, Lindley DA, Patel N, et al. Twelve-month analgesia and rescue, by cooled radiofrequency ablation treatment of osteoarthritic knee pain: results from a prospective, multicenter, randomized, cross-over trial. Reg Anesth Pain Med. 2019:rapm-2018-100051. doi: 10.1136/rapm-2018-100051.
9. Hunter C, Davis T, Loudermilk E, Kapural L, DePalma M. Cooled Radiofrequency Ablation Treatment of the Genicular Nerves in the Treatment of Osteoarthritic Knee Pain: 18- and 24-Month Results. Pain Pract. 2020;20(3):238-46. doi: 10.1111/papr.12844.
10. Jüni P, Hari R, Rutjes AW, Fischer R, Silletta MG, Reichenbach S, et al. Intra-articular corticosteroid for knee osteoarthritis. Cochrane Database Syst Rev. 2015;2015(10):CD005328. doi: 10.1002/14651858.CD005328.pub3.
11. Hong T, Wang H, Li G, Yao P, Ding Y. Systematic Review and Meta-Analysis of 12 Randomized Controlled Trials Evaluating the Efficacy of Invasive Radiofrequency Treatment for Knee Pain and Function. Biomed Res Int. 2019;2019:9037510. doi: 10.1155/2019/9037510.
12. Li G, Zhang Y, Tian L, Pan J. Radiofrequency ablation reduces pain for knee osteoarthritis: A meta-analysis of randomized controlled trials. Int J Surg. 2021;91:105951. doi: 10.1016/j.ijsu.2021.105951.

13. Zhang H, Wang B, He J, Du Z. Efficacy and safety of radiofrequency ablation for treatment of knee osteoarthritis: a meta-analysis of randomized controlled trials. J Int Med Res. 2021;49(4):3000605211006647. doi: 10.1177/03000605211006647.
14. Kim DH, Lee MS, Lee S, Yoon SH, Shin JW, Choi SS. A Prospective Randomized Comparison of the Efficacy of Ultrasound- vs Fluoroscopy Guided Genicular Nerve Block for Chronic Knee Osteoarthritis. Pain Physician. 2019;22(2):139-46.
15. Hmamouchi I, Allali F, Tahiri L, Khazzani H, El Mansouri L, Alla SAO, et al. Clinically important improvement in the WOMAC and predictor factors for response to non-specific non-steroidal antiinflammatory drugs in osteoarthritic patients: a prospective study. BMC Res Notes. 2012;5:58. doi: 10.1186/1756-0500-5-58.
16. Lizaur-Utrilla A, Gonzalez-Parreño S, Martinez-Mendez D, Miralles-Muñoz FA, Lopez-Prats FA. Minimal clinically important differences and substantial clinical benefits for Knee Society Scores. Knee Surg Sports Traumatol Arthrosc. 2020;28(5):1473-8. doi: 10.1007/s00167-019-05543-x.
17. Konya ZY, Akin Takmaz S, Başar H, Baltaci B, Babaoğlu G. Results of genicular nerve ablation by radiofrequency in osteoarthritis-related chronic refractory knee pain. Turk J Med Sci. 2020;50(1):86-95. doi: 10.3906/sag-1906-91.
18. Martins GC, Camanho GL, Ayres LM, Oliveiras ES. Correlation between Ahlback radiographic classification and anterior cruciate ligament status in primary knee arthrosis. Rev Bras Ortop. 2017;52(1):69-74. doi: 10.1016/j.rboe.2016.02.012.
19. McCormick ZL, Reddy R, Korn M, Dayanim D, Syed RH, Bhave M, et al. A Prospective Randomized Trial of Prognostic Genicular Nerve Blocks to Determine the Predictive Value for the Outcome of Cooled Radiofrequency Ablation for Chronic Knee Pain Due to Osteoarthritis. Pain Med. 2018;19(8):1628-38. doi: 10.1093/pm/pnx286.
20. Conger A, Gililland J, Anderson L, Pelt CE, Peters C, McCormick ZL. Genicular Nerve Radiofrequency Ablation for the Treatment of Painful Knee Osteoarthritis: Current Evidence and Future Directions. Pain Med. 2021;22(Suppl 1):S20-3. doi: 10.1093/pm/pnab129.
21. Iannaccone F, Dixon S, Kaufman A. A Review of Long-Term Pain Relief after Genicular Nerve Radiofrequency Ablation in Chronic Knee Osteoarthritis. Pain Physician. 2017;20(3):E437-44.
22. Kapural L, Minerali A, Sanders M, Matea M, Dua S. Cooled Radiofrequency Ablation Provides Prolonged Pain Relief Compared to Traditional Radiofrequency Ablation: A Real-World, Large Retrospective Clinical Comparison from a Single Practice. J Pain Res. 2022;15:2577-86. doi: 10.2147/JPR.S373877.
23. Tran A, Gonzalez FM. Review of cooled radiofrequency ablation utilization for the treatment of symptomatic advanced knee arthritis and total knee arthroplasty. Skeletal Radiol. 2023;52(5):941-9. doi: 10.1007/s00256-022-04058-w.
24. Chou SH, Shen PC, Lu CC, Liu ZM, Tien YC, Huang PJ, et al. Comparison of Efficacy among Three Radiofrequency Ablation Techniques for Treating Knee Osteoarthritis: A Systematic Review and Meta-Analysis. Int J Environ Res Public Health. 2021;18(14):7424. doi: 10.3390/ijerph18147424.
25. Chen Y, Vu TH, Chinchilli VM, Farrag M, Roybal AR, Huh A, et al. Clinical and technical factors associated with knee radiofrequency ablation outcomes: a multicenter analysis. Reg Anesth Pain Med. 2021;46(4):298-304. doi: 10.1136/rapm-2020-102017.
26. Santana-Pineda M, Vaulinthout L, Santana-Ramírez S, Vanneste T, Zundert JV, Novalbos-Ruiz JP. A randomized controlled trial to compare analgesia and functional improvement after continuous neuroblative and pulsed neuromodulated radiofrequency. Pain Med. 2021;22(3):637-52. doi: 10.1093/pm/pnaa309.
27. Caragea M, Woodworth T, Curtis T, Blatt M, Cheney C, Brown T, et al. Genicular nerve radiofrequency ablation for the treatment of chronic knee joint pain: a real-world cohor study with evaluation of prognostic factors. Pain Med. 2023;24(12):1332-40. doi: 10.1093/pm/pnad095.
28. Dias J, Moraes V, Freitas M, Arouca MM, Ocampos GP, Camargo OP, et al. Genicular nerve ablation in knee osteoarthritis: A randomized prospective study. Acta Ortop Bras. 2025,33(3):e289218. doi: 10.1590/1413-785220253303e289218.

DOI: http://dx.doi.org/10.1590/1413-785220263404e304880

TREATMENT OF BENIGN LYTIC BONE LESIONS WITH 45S5 BIOACTIVE GLASS: A CASE SERIES

TRATAMENTO DE LESÕES ÓSSEAS LÍTICAS BENIGNAS COM BIOVIDRO 45S5: SÉRIE DE CASOS

Jaime Gomes Nogueira[1], *Marco de Andrade Bianchi*[1], *André Ferrari de França Camargo*[1], *André Mathias Baptista*[1], *Olavo Pires de Camargo*[1]

1. Universidade de Sao Paulo, Faculdade de Medicina, Hospital das Clinicas, Instituto de Ortopedia e Traumatologia IOT-HCFMUSP, Sao Paulo, SP, Brazil.

ABSTRACT

Objective: To report a series of 9 cases of patients with benign cavitary tumor lesions who underwent intralesional resection (curettage) followed by cavity filling with 45S5 bioactive glass (Aktibone). Methods: This is a therapeutic case series study with level IV evidence. Patients diagnosed with benign lytic bone lesions who had a surgical indication for curettage followed by cavity filling with bioactive glass were included. In the postoperative period, patients underwent imaging examinations, which were qualitatively evaluated by a radiologist who determined the radiological diagnosis and the quality of osseointegration. The presence of imaging abnormalities was also assessed, and statistical analysis was performed. Results: Nine patients were included, 7 males and 2 females. The mean age was 25.8 years. The initial diagnoses varied, with enchondromas being the most frequent. The mean follow-up time after surgery was 8.4 months. All patients underwent preoperative radiographs. Most patients showed good clinical progression and good to partial osseointegration. Conclusion: The results reinforce that 45S5 bioactive glass demonstrates good performance as a bone substitute in benign lytic lesions, providing adequate osseointegration and a low complication rate. ***Level of Evidence IV; Case series.***

Keywords: Bone Neoplasms; Osseointegration; Biocompatible Materials; Orthopedic Surgery.

RESUMO

Objetivo: Relatar uma série de 9 casos de pacientes com lesões tumorais cavitárias benignas submetidos a ressecção intralesional (curetagem) e posterior preenchimento com biovidro 45S5 (Aktibone). Métodos: É um estudo terapêutico do tipo série de casos com nível de evidência IV. Foram incluídos pacientes com diagnóstico de lesão óssea lítica benigna e que tinham indicação cirúrgica de curetagem da lesão e posterior preenchimento da cavidade com o biovidro. No pós-operatório, os pacientes realizaram exames de imagem que foram avaliados de forma qualitativa por um médico radiologista que definiu o diagnóstico radiológico e a qualidade da osseointegração. Também foi observada a presença de anormalidades na imagem e foi realizada análise estatística. Resultados: Foram incluídos 9 pacientes, sendo 7 homens e 2 mulheres. A média de idade foi 25,8 anos. Os diagnósticos iniciais foram variados, com predominância de encondromas. O tempo de acompanhamento pelo serviço após a realização da cirurgia teve média de 8,4 meses. Todos os pacientes realizaram RX pré-operatório. A maioria dos pacientes tiveram boa evolução e osteointegração boa e parcial. Conclusão: Os resultados obtidos reforçam que o biovidro 45S5 apresenta bom desempenho como substituto ósseo em lesões líticas benignas, proporcionando adequada osteointegração e baixo índice de complicações. ***Nível de Evidência IV; Série de casos.***

Descritores: *Neoplasias Ósseas; Osseointegração; Materiais Biocompatíveis; Cirurgia Ortopédica.*

Citation: Nogueira JG, Bianchi MA, Camargo AFF, Baptista AM, Camargo OP. Treatment of benign lytic bone lesions with 45S5 bioactive glass: a case series. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 4. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

The prevalence of benign bone lesions of low aggressiveness remains poorly documented, mainly due to their often asymptomatic presentation and insidious evolution, often being detected incidentally during complementary exams requested for the investigation of fractures or other musculoskeletal conditions.[1] Among the main etiologies, enchondroma, osteoma, non-ossifying fibroma, and unicameral bone cysts stand out, which, although benign, may require surgical treatment to prevent pathological fractures and restore function.[1,2]

The most commonly employed surgical approach in these cases is curettage, whose objective is the complete removal of the pathological tissue while preserving as much healthy bone as possible.[3] Subsequently, filling the resulting cavity is a fundamental step

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Instituto de Ortopedia e Traumatologia HCFMUSP
Correspondence: André Mathias Baptista. Rua Dr. Ovídio Pires de Campos, 333, São Paulo, SP, Brazil, 05403010. andremathiasbaptista@gmail.com

Article received on 01/23/2026 approved on 03/10/2026.
Handling Editor: Andre Ferrari de Camargo

to restore mechanical strength and promote bone regeneration. Thus, autologous bone grafts are traditionally considered the gold standard due to their excellent osteogenicity, osteoinductivity, and osteoconductivity; however, they have significant limitations, such as morbidity of the donor site, pain, and limited availability.[3,4] Alternatively, allografts are widely used, but they carry risks of pathogen transmission, immune reactions, and biomechanical failures. Other options include xenografts, bone cement, and synthetic bone substitutes, which vary in biological properties, mechanical properties, and cost-effectiveness.[3]

In light of this, among synthetic substitutes, the development of bioactive glasses marked a milestone in biomaterials engineering, being the first compounds capable of establishing a direct chemical bond with bone tissue.[5] The bioglass exhibits controlled degradation and the release of bioactive ions, such as soluble silica, calcium, sodium, and potassium, which stimulate cell proliferation, bone neoformation, and angiogenesis, through the crystallization of apatite on its surface and modulation of the local microenvironment. The classic formulation 45S5, commercially known as Bioglass®, composed of 46.1% SiO_2, 26.9% CaO, 24.4% Na_2O, and 2.6% P_2O_5, demonstrates strong integration with the bone matrix, such that its removal is unfeasible without fracturing the adjacent bone after complete remodeling, which occurs between 6 and 18 months post-operation.[2,5-7]

In addition to its osteoconductive and osteoinductive properties, active bioglass exhibits an intrinsic antimicrobial effect, related to the increase in pH and osmolarity at the implanted site, making it unfavorable for bacterial growth.[5] It is a safe, non-toxic, and chemically stable biomaterial that combines regenerative potential with infection prevention, positioning itself as a promising alternative in the management of low-aggressiveness lytic bone lesions.[6,8]

In this context, recent comparative trials and randomized studies reinforce that bioactive glass 45S5 graft substitutes provide outcomes similar to those of allografts in terms of lesion recurrence, need for reoperation, and functional scores in heterogeneous series of benign tumors.[2,9,10] However, there are limitations in the global literature regarding factors such as lesion heterogeneity and relatively short follow-up, which impose caution in generalization, in addition to highlighting the need for multicenter studies with prolonged follow-up and long-term remodeling assessment.[2]

In this context, the evaluation of active bioglass 45S5 in bone lesions through imaging exams is essential to monitor the incorporation of the material, bone neoformation, and structural stability over time. Computed tomography (CT) has been widely used to quantify bone mineral density and monitor graft remodeling, allowing differentiation between residual material and regenerated bone.[10] Magnetic resonance imaging (MRI), in turn, provides information about tissue integration and potential complications, such as tumor recurrence or inflammatory reaction, without exposure to ionizing radiation.[9] Furthermore, recent studies highlight the role of serial radiography as an accessible and effective method for monitoring clinical evolution, although with sensitivity limitations compared to CT.[2] Therefore, the combination of different imaging modalities has been recommended to provide a more comprehensive analysis of osteointegration and the safety of bioglass 45S5, ensuring greater diagnostic accuracy and support for clinical decision-making.[6] Thus, the objective of this study is to report a series of 9 cases of patients with low-aggressiveness cavitary tumor lesions who underwent intralesional resection (curettage) and subsequent filling with bioglass 45S5 (Aktibone).

MATERIALS AND METHODS

This is an observational, descriptive, and prospective study of case series. This is a therapeutic study with a level of evidence IV. The present research was conducted following the ethical standards established by the Nuremberg Code (1947), the Declaration of Helsinki (2000), and the Research Standards Involving Human Beings established by Resolution 466/12 of the National Health Council. This study was approved by the Research Ethics Committee (CEP) of IOT-HC-FMUSP, under opinion CAAE 58963622.3.0000.0068. The evaluated patients signed the Informed Consent Form (ICF).

The work was conducted by the Orthopedic Oncology Group of the Institute of Orthopedics and Traumatology of the Hospital das Clínicas of the Universidade de São Paulo (IOT-HCFMUSP). The surgeries were performed at the Surgical Center of the Institution.

MATERIALS

The bioactive glass granules (size between 500-1000 μm), commercially named Aktibone, produced by Noraker Bioglass Company, Lyon, France, were fully donated by RCL Implantes, the official distributor of the product in Brazil, imported by Visão Implantes. The material, designated as bioglass 45S5, is composed of 45% SiO2, 24.5% Na2O, 24.5% CaO, and 6% P2O5.

Inclusion criteria

Patients enrolled at IOT-HCFMUSP with a diagnosis of low-aggressiveness lytic bone lesions, such as enchondromas, non-ossifying fibromas, and simple bone cysts, were included. Furthermore, the patients had surgical indications for various reasons, such as risk of fracture or orthopedic complications, pain, an increase in the size of the lesion, and risk of oncological progression.

Data collection

The selected patients followed the usual routine flow of the Orthopedic Oncology Group.

The surgeries were performed according to good medical practice and with the established surgical techniques already in use by the Orthopedic Oncology Group. The bone lesions were accessed through a bone window created in the cortex using chisels; after curettage of the lesion, the internal walls of the lesion were smoothed with a high-speed drill, thus reducing the risk of residual neoplastic cells remaining in small recesses or spaces created by the microtrabecular bone; after this smoothing, the internal walls were cauterized with an electric scalpel. These measures are important for reducing the risk of local recurrence. Next, the cavity was filled with bioglass, and, whenever possible, the removed bone window was replaced to occlude the opening. The decision to add a prophylactic metal implant (plate and screws) depended on the anatomical region and dimensions of the lesion, and was made on a case-by-case basis.

The immediate postoperative procedures were performed as routinely done by the Group (prophylactic antibiotic therapy for 24 hours, dressing care, drains, motor physiotherapy, etc.). After discharge, the patients returned at intervals of 1, 2, and 4 weeks for evaluation of soft tissue healing and any acute complications; and at 3, 6, and 12 months for evaluation of consolidation, osteointegration, functional recovery, and any late complications. For this evaluation, imaging tests were used, such as X-ray, computed tomography, and magnetic resonance imaging.

Data analysis and statistics

In the postoperative evaluation, the patients underwent imaging tests that were qualitatively assessed by a radiologist from the service. After a summative analysis of all the tests performed by each patient, the radiological diagnosis was defined as: good evolution, poor evolution, or inconclusive. Additionally, regarding osseointegration, it could be classified as: good and total, good and partial, moderate, poor, and inconclusive. Abnormalities such

as osteolysis, expansive calcifications, deformities, osteoporosis, and fractures, among others, were also observed.
Descriptive statistics of the quantitative data were performed: mean, median, maximum, and minimum values for each parameter.

Risk analysis

Considering that bioglass 45S5 (Aktibone) is already a product widely used in the Brazilian market in various clinical situations) with no reports of complications, the risk of this research was considered low.

RESULTS

Nineteen eligible patients were selected according to the inclusion criteria; however, ten were excluded during the research, with one patient due to a postoperative complication of fracture and loss of bone alignment, while the others (nine) lost follow-up after surgery. Thus, nine patients with low-aggressiveness lytic bone disease were included in the study, consisting of seven men and two women. Regarding age, the youngest patient was seven years old, and the oldest was fifty-four years old, with an average age of 25.8 years and a median of 22 years. The initial diagnosis was varied and included: Six enchondromas, two non-ossifying fibromas, and one simple bone cyst. The location of the lesions was diverse and is better described in Table 1.

Table 1. Location and diagnosis of bone lesions.

Patient	Diagnosis	Location of the lesion
A	Non-ossifying fibroma	Right femur
B	Enchondroma	Proximal phalanx of the right fifth digit
C	Enchondroma	Right femur
D	Simple bone cyst	Left humerus
E	Enchondroma	Right third metacarpal
F	Enchondroma	Distal phalanx of the right first digit
G	Non-ossifying fibroma	Right distal tibia
H	Enchondroma	Middle phalanx of the right fourth digit
I	Enchondroma	Proximal phalanx of the right fifth digit

Source: author, 2025.

Regarding the follow-up time by the service after the surgery, the average was 8.4 months, and the median was 8 months, with a minimum of 2 months and a maximum of 13 months. All patients underwent pre-operative X-rays. In the post-operative period, the average number of imaging exams performed by patients was 4.7 and the median was 5, with the minimum value corresponding to the patient who underwent 2 exams and the maximum to the patient who underwent 7 exams. The radiologist made qualitative observations of the lesions, which are better described in Table 2. Thus, it is observed that the majority of patients (6) had good evolution and good and partial osteointegration (3).

DISCUSSION

The present study evaluated the use of bioglass 45S5 in the treatment of low-aggressiveness lytic bone lesions, including nine patients, predominantly male (77.7%), with an average age of 25.8 years. The lesions were classified as benign or non-aggressive behavior, including enchondromas, non-ossifying fibromas, and simple bone cysts. The majority of patients (66.7%) showed good clinical and radiographic evolution, with signs of satisfactory osteointegration over an average follow-up of 8.4 months, demonstrating the osteoconductive and osteoinductive potential of bioglass 45S5 in cavitary bone defects.
These findings are consistent with recent evidence from the literature confirming the potential of bioglass 45S5 as an effective and safe bone substitute. Nogueira et al. (2024) highlight that 45S5 exhibits high bioactivity, with ionic release of calcium, phosphorus, and silicon, promoting the formation of a carbonated hydroxyapatite layer that favors osteointegration.[6] This characteristic stimulates osteoblastic differentiation and accelerates the bone regeneration process, making the material a promising alternative to autologous grafts, which still represent the gold standard but with limitations associated with donor site morbidity.[6]
Moreover, clinical studies also corroborate the good response of bioglass in benign bone lesions. Thus, Samade et al. (2022), in a series of pediatric cases, observed good incorporation of the material and absence of major complications in cavity defects

Table 2. Qualitative and summative assessment of radiological imaging exams..

Patient	Evolution	Osteointegration	Number of imaging exams performed	Types of imaging exams performed	Follow-up time	Observations
A	Inconclusive	Inconclusive	2	RX	2 months	Presence of expansive calcifications
B	Poor	Moderate	7	RX, TC, RM	13 months	Presence of osteolysis in about 70% of the area + fracture and deformity
C	Good	Good and partial	7	RX, TC, RM	13 months	—
D	Good	Good and total	5	RX, TC	8 months	Evolution with osteoporosis of the limb
E	Poor	Poor	5	RX, TC, RM	13 months	Progression of cartilage injury + fracture + Presence of a halo of osteolysis around the bioglass + No closure of the dorsal cortex
F	Good	Good and total	2	RX	2 months	Presence of a thin halo of osteolysis in the distal portion
G	Good	Moderate	5	RX, TC, RM	8 months	Presence of a halo of osteolysis around the bioglass + No integration with the adjacent bone marrow + No closure of the anterior cortex
H	Good	Good and partial	7	RX, TC, RM	11 months	Evolution with re-ossification
I	Good	Good and partial	3	RX	6 months	Presence of partial resorption on the medial face of the middle third

Source: author, 2025.

treated with bioglass 45S5 after curettage of benign bone tumors.[11] The average consolidation time was similar to that observed with autologous grafts, with a lower risk of recurrence and no need for surgical reintervention. Similarly, Ma et al. (2021), when comparing the use of bioglass 45S5 with allogenic bone grafts in symptomatic lesions of the calcaneus, reported equivalent clinical results, with radiographic evidence of progressive bone formation and good integration of the biomaterial.[10]

In terms of clinical applicability, bioglass 45S5 presents significant advantages, such as excellent biocompatibility, absence of immunogenic response, and ease of intraoperative molding. Recent experimental studies also demonstrate that the incorporation of therapeutic ions, such as copper or strontium, can enhance the osteogenic and antimicrobial effects of 45S5, expanding its clinical application possibilities.[12]

In the present study, the presence of three cases with partial osteointegration of enchondromas may be related to the location of the lesions, in the femur and phalanges, or to the volume of the treated defect, factors already pointed out by Jin, Neuville, and Brauer (2025) as determinants in the biomechanical performance of bioglasses, since the amount of 45S5 applied directly influences the regeneration process and the mechanical strength of the newly formed bone.[12] Additionally, individual differences in bone metabolism, age, and quality of the receptor bed may also interfere with the biological response to the biomaterial.[12]

Another relevant aspect is the short average follow-up time (8.4 months) observed in this study. Although the initial results are encouraging, the literature indicates that complete replacement of bioglass with mature bone tissue may take up to 18 to 24 months.[6,11] Thus, prolonged follow-up is essential to confirm structural stability and absence of recurrence of the lesions, especially in tumors with more active behavior. Such evolution, however, is hindered by the loss of follow-up of patients in the orthopedic service.

Regarding the choice of the best imaging exam to evaluate the osteointegration of bioglass in the postoperative period, it was noted that computed tomography is more suitable as it provides better visualization of the lesion, in accordance with Tsukayama et al. (1999).[4] However, due to the more difficult access to CT and the higher radiation used, serial X-rays may be used for monitoring in cases of good evolution and good osteointegration, opting for CT only in complicated cases or those that do not show good evolution on X-ray. Such analysis is consistent with the study by Incesoy et al. (2025), in which radiographs were performed at different times (6 weeks, 3, 6, 12, and 24 months) to monitor bone filling in the postoperative period and with Giavaresi et al. (2008), in their randomized prospective study with 12 patients where imaging studies were performed with serial simple radiographs, in addition to computed tomography, magnetic resonance imaging, bone scintigraphy, and SPECT.[2,13]

On the other hand, it should be considered that the small sample size and the absence of a control group limit the extrapolation of the results. The diagnostic and anatomical heterogeneity of the lesions also constitutes a bias factor, making it difficult to standardize the response to treatment. Despite these limitations, the findings are consistent with recent literature and suggest that bioglass 45S5 represents a safe and effective alternative for the treatment of bone defects resulting from benign lytic lesions, especially in young patients with good bone quality.

CONCLUSION

Therefore, the results obtained in this series of cases reinforce that bioglass 45S5 performs well as a bone substitute in low-aggressiveness lytic lesions, providing adequate osteointegration and a low complication rate. However, studies with larger samples, longer follow-up, and objective analysis of bone regeneration are necessary to consolidate its role in the surgical management of these lesions.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. COP: substantial contribution to the study design and performance of surgeries; BAM: substantial contribution to the interpretation of the study data and final approval of the manuscript version to be published; CAFF: data acquisition and writing; NJG: data acquisition, performance of surgeries, and writing; BMA: interpretation of the data for the study and writing.

DATA AVAILABILITY DECLARATION

The underlying contents of the research text are contained in the manuscript.

REFERENCES

1. Conde RA, Silva Oliveira Filho AR, Machinski E, da Cruz VF, Melinte MA, Varone BB, et al. Can bioactive glass replace bone grafts after curettage of benign bone tumors? A systematic review and meta-analysis of randomized controlled trials. J Orthop. 2025;69:208-15. doi: 10.1016/j.jor.2025.05.061.
2. Incesoy MA, Sarikas M, Kaya HB, Yavuz AD, Sulak B, Ozdemir AA, et al. Bioactive glass graft versus allograft in benign bone lesions: a retrospective comparative study. HSS J. 2025;21(4):396-403. doi: 10.1177/15563316251321825.
3. Gava NF, Engel EE. Treatment alternatives and clinical outcomes of bone filling after benign tumor curettage: a systematic review. Orthop Traumatol Surg Res. 2022;108(4):102966. doi: 10.1016/j.otsr.2021.102966.
4. Tsukayama DT. Pathophysiology of posttraumatic osteomyelitis. Clin Orthop Relat Res. 1999;(360):22-9. doi: 10.1097/00003086-199903000-00005.
5. Dash PA, Mohanty S, Nayak SK. A review on bioactive glass, its modifications and applications in healthcare sectors. J Non-Cryst Solids. 2023;605:122404. doi: 10.1016/j.jnoncrysol.2023.122404.
6. Nogueira DMB, Rosso MPO, Buchaim DV, Zangrando MSR, Buchaim RL. Update on the use of 45S5 bioactive glass in the treatment of bone defects in regenerative medicine. World J Orthop. 2024;15(3):204-14. doi: 10.5312/wjo.v15.i3.204.
7. Cannio M, Bellucci D, Roether JA, Boccaccini DN, Cannillo V. Bioactive glass applications: a literature review of human clinical trials. Materials. 2021;14(18):5440. doi: 10.3390/ma14185440.
8. Camargo AFF, Baptista AM, Natalino R, Camargo OP. Bioactive glass in cavitary bone defects: a comparative experimental study in rabbits. Acta Ortop Bras. 2015;23(4):202-7. doi: 10.1590/1413-785220152304147538.
9. Syvänen J, Serlo W, Jalkanen J, Kahonen I, Raitio A, Nietosvaara Y, et al. Allograft versus bioactive glass (BG-S53P4) in pediatric benign bone lesions: a randomized clinical trial. J Bone Joint Surg Am. 2023;105(9):659-66. doi: 10.2106/JBJS.22.00716.
10. Ma H, Shi Y, Zhang W, Liu F, Han Y, Yang M. Open curettage with bone augmentation for symptomatic tumors and tumor-like lesions of calcaneus: a comparison of bioactive glass versus allogeneic bone. J Foot Ankle Surg. 2021;60(5):881-6. doi: 10.1053/j.jfas.2021.02.010.
11. Samade R, Jonard B, Scharschmidt TJ. Utility of 45S5 bioactive glass as bone graft substitute for cavitary defects in pediatric bone tumors: a preliminary study. J Pediatr Orthop. 2022;42(9):532-8. doi: 10.1097/BPO.0000000000002232.
12. Jin Z, Neuville DR, Brauer DS. Glasses for bone regeneration: structural features controlling physical properties and ion release of bioactive glasses 45S5, S53P4 and 13-93. RSC Adv. 2025;15:4997-5006. doi: 10.1039/D4RA06081D.
13. Giavaresi G, Borsari V, Fini M, Giardino R, Sambri V, Gaibani P, et al. Preliminary investigations on a new gentamicin and vancomycin-coated PMMA nail for the treatment of bone and intramedullary infections: An experimental study in the rabbit. J Orthop Res. 2008;26(6):785-92. doi: 10.1002/jor.20543.

DOI: http://dx.doi.org/10.1590/1413-785220263404e300987

SCREW VS. SUTURE FIXATION FOR TIBIAL EMINENCE FRACTURES: A META-ANALYSIS

PARAFUSO VS. SUTURA NA FIXAÇÃO DE FRATURAS DA ESPINHA TIBIAL: UMA META-ANÁLISE

Paulo Eduardo Dias Lavigne[1], *Marcus Vinicius Ribeiro Filho*[1], *Victor Mendes Ribeiro*[1], *Jessiel Gonçalves de Souza*[1], *Adriano Assis Nascimento*[1], *Eiffel Tsuyoshi Dobashi*[1]

1. Universidade Federal de Sao Paulo (EPM/UNIFESP), Escola Paulista de Medicina, Sao Paulo, SP, Brazil.

ABSTRACT

Objectives: To compare functional and surgical outcomes between screw and suture fixation in pediatric tibial spine fractures. Materials and Methods: This systematic review followed PRISMA guidelines and was registered in PROSPERO (CRD420251022233). Cohort studies comparing arthroscopic screw versus suture fixation were included. The primary outcome assessed was postoperative knee function. Results: Four retrospective cohort studies were included, totaling 186 patients. No significant differences were found in knee function (SMD = –0.07; p = 0.702). Operative time was significantly shorter in the screw group (MD = –9.17 minutes; p < 0.001). However, this group showed a significantly higher risk of reoperation (RR = 1.78; p = 0.040) and implant removal (RR = 6.25; p = 0.040). No differences were observed regarding joint instability (RR = 0.90; p = 0.770) or return to sport (RR = 1.83; p = 0.190). Funnel plot analysis suggested potential publication bias. Conclusion: Both techniques yielded similar postoperative functional outcomes. Screw fixation was associated with shorter operative time but higher complication rates. ***Level of evidence IIa; systematic review of cohort studies.***

Keywords: Tibial Spine Fractures; Anterior Cruciate Ligament Injuries; Orthopedic Surgery; Arthroscopy; Internal Fixators; Meta-Analysis as Topic; Knee Joint.

RESUMO

Objetivos: Comparar os desfechos funcionais e cirúrgicos entre parafuso e sutura na fixação de fraturas de espinha tibial em pacientes pediátricos. Material e Métodos: Esta revisão sistemática seguiu o guideline PRISMA, registrada no PROSPERO (CRD420251022233). Foram incluídos estudos de coorte que compararam, por via artroscópica, a fixação com parafuso versus sutura. O desfecho primário analisado foi verificar a função do joelho no período pós-operatório. Resultados: Foram incluídos quatro estudos de coorte retrospectivos, totalizando 186 pacientes. Não foram observadas diferenças significativas na função do joelho (DMP = –0,07; p = 0,702). O tempo operatório foi significativamente menor no grupo tratado com parafuso (DM = –9,17 minutos; p < 0,001). No entanto, esse grupo apresentou risco significativamente maior de reoperação (RR = 1,78; p = 0,040) e de remoção do implante (RR = 6,25; p = 0,040). Não foram encontradas diferenças em relação à instabilidade articular (RR = 0,90; p = 0,770) ou ao retorno à prática esportiva (RR = 1,83; p = 0,190). A análise do gráfico de funil sugeriu possível viés de publicação. Conclusão: Ambas as técnicas apresentaram resultados funcionais pós-operatórios semelhantes. A fixação com parafuso foi associada a menor tempo cirúrgico, porém com maiores taxas de complicações. ***Nível de evidência IIa; revisão sistemática de estudos de coorte.***

Descritores: *Fraturas da Espinha Tibial; Lesões do Ligamento Cruzado Anterior; Cirurgia Ortopédica; Artroscopia; Fixadores Internos; Meta-Análise como Assunto; Articulação do Joelho.*

Citation: Lavigne PED, Ribeiro Filho MV, Ribeiro VM, Souza JG, Nascimento AA, Dobashi ET. Screw vs. Suture fixation for tibial eminence fractures: a meta-analysis. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 7. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

Tibial spine fractures are caused by the avulsion of the anterior cruciate ligament (ACL) at its insertion on the intercondylar eminence, occurring more frequently in children. Although uncommon, such traumatic disorders have clinical relevance and require careful evaluation, especially in cases with displacement.[1,2] The Meyers and McKeever classification remains the most commonly used reference to guide surgical indication in type II fractures with displacement, as well as in type III and IV fractures.[3]

The most commonly employed fixation techniques include the use of metal screws (SW) and non-absorbable suture fixation (SF), both frequently performed arthroscopically.[4] These approaches aim to restore the anatomical alignment of the tibial spine, preserve ACL function, and enable a safe return to sports activities.[5] The choice

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Universidade Federal de Sao Paulo (EPM/UNIFESP), Escola Paulista de Medicina, Sao Paulo, SP, Brazil.
Correspondence: Paulo Eduardo Dias Lavigne. 740, R. Botucatu, Vila Clementino, Sao Paulo, SP, Brazil. 04023-062.paulo_lavigne2@hotmail.com

Article received on 16/09/2025 approved on 22/09/2025.
Handling Editor: Claudio Santilli

of technique depends on the fracture morphology, the patient's age, the desired degree of stability, and the surgeon's experience. However, uncertainty persists regarding the ideal surgical strategy.[6] Biomechanical studies have reported conflicting results regarding mechanical strength between different fixation methods, and direct clinical evidence remains limited.[7,8] Although both techniques have demonstrated satisfactory outcomes in clinical practice, it is still unclear whether one presents superiority in essential aspects such as knee function, complication rates, operative time, or return to sport.[9,10] In the absence of randomized clinical trials (RCTs) and given the predominance of observational studies, this systematic review with meta-analysis and meta-regression aims to synthesize comparative observational data on screw versus suture fixation in tibial spine fractures. By analyzing functional and surgical outcomes in different clinical contexts, this study seeks to enhance the understanding of the patterns of results observed in practice, without the intention of establishing definitive recommendations, but rather to support a more informed surgical decision-making process.

MATERIALS AND METHODS

This systematic literature review was conducted according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[11]. The primary objective was to compare the clinical and functional outcomes of screw fixation versus arthroscopic suture fixation in children and adolescents with tibial spine fractures.

The PICOTT strategy was applied as follows: Population (P) – Children and adolescents with tibial spine fractures/tibial eminence fractures/anterior cruciate ligament (ACL) avulsion fractures treated surgically; Intervention (I) – Screw fixation; Comparison (C) – Arthroscopic suture fixation; Outcomes (O) – Primary: Knee function and range of motion (ROM). Secondary outcomes: Return to sport, knee stability, postoperative complications, reoperation rates, time to radiographic bone consolidation; Study type (T) – Randomized clinical trials (RCTs) and cohort studies (prospective or retrospective); Time (T) – Any follow-up time.

This review was registered in the International Prospective Register of Systematic Reviews (PROSPERO) under registration number CRD420251022233. The methodology was defined and registered prior to the start of the search in the databases.

Eligibility criteria

Eligible studies included full-text publications of RCTs, written in English and available until April 2025 in the PubMed, Scopus, Embase, and Cochrane CENTRAL databases. The studies should compare screw fixation and arthroscopic suture fixation, reporting clinical outcomes related to functional recovery and complications. Articles not available in full text through digital platforms, conference abstracts, preprints (*preprints)*, letters to the editor, case reports and series, cross-sectional studies, and case-control studies, reviews, and studies without sufficient data on relevant outcomes were excluded. Duplicate publications of the same study were excluded, retaining only the most complete version.

Search strategy

To search for articles in the databases, descriptors related to "Tibial Spine Fractures," "Screw Fixation," and "Arthroscopic Suture" were used. The descriptors were obtained from *Medical Subject Headings* (MeSH), accessed at www.ncbi.nlm.nih.gov/mesh/. Boolean operators AND and OR were employed to combine the terms on the mentioned platforms, respecting the inclusion and exclusion criteria of the articles. All descriptors used, as well as the complete search strategy for each database, are available in Supplementary Material 1 and 2. [12,13]

Study selection

Two independent reviewers, blinded to each other's assessments, jointly screened the titles and abstracts of all retrieved articles to identify those that met the previously established inclusion criteria. Potentially eligible studies were then read in full to confirm inclusion. In case of disagreement, the final decision was made by a senior reviewer, who had access only to the articles related to the research. The selection of studies was conducted using the Rayyan application.[14] To ensure the comprehensiveness of the available literature, the snowballing strategy was also used:[15] the reference lists of all relevant systematic reviews identified in the initial search were examined to identify additional eligible primary studies, and the references of the full-text articles were also analyzed after full reading to detect studies possibly not captured in the original search.

Data extraction

Data extraction was performed independently and in duplicate (PEDL and JGS) to ensure the accuracy and reliability of the information. The data from the included studies were extracted using a pre-prepared form in Microsoft Excel® (version 2205), containing comprehensive information about the characteristics of the studies, including sample size, details of the intervention and control groups, methodological aspects, and the outcomes assessed. Any discrepancies between the reviewers were resolved by consensus or with the mediation of a senior reviewer (ETD).

Quality Assessment

The assessment of bias risk was conducted based on the *Risk Of Bias In Non-randomized Studies - of Interventions* (ROBINS-I) tool proposed by Cochrane,[16] applied independently by two reviewers (PEDL and JGS), with discrepancies resolved by consensus. Publication bias was assessed using a funnel plot with contour of statistical significance (*contour-enhanced funnel plot*), which displays point estimates based on the weight of each study and highlights areas of statistical significance to facilitate interpretation.[17]

Statistical Analysis

The primary outcome of this meta-analysis was to verify knee function in the postoperative period. Given the use of distinct validated instruments among the studies (IKDC Subjective Scale and Lysholm score), this outcome was evaluated using the standardized mean difference (SMD), with 95% confidence intervals (CI95%). For operative time, uniformly measured in minutes, the mean difference (MD) was used. Dichotomous outcomes (such as knee instability in the postoperative period, failure to return to sport, reoperations, and implant removal) were analyzed using relative risk (RR), also with CI95%.

Heterogeneity was assessed with the Cochran Q test and the I^2 index, with p values < 0.10 and $I^2 > 25\%$ considered indicative of substantial heterogeneity. Prediction intervals (PI) were also reported to express the expected variation in effect sizes in similar future studies. Given the clinical and methodological variability among the included studies, all meta-analyses were conducted using a random effects model, applying the Restricted Maximum Likelihood (REML) estimator.

For the specific outcome of implant removal, due to the low frequency of these events and the limited number of studies, methodological adjustments were made to increase the robustness of the estimates: (1) continuity correction, by adding 0.5 to all cells of the contingency tables to account for data with zero events; and (2) confidence intervals were calculated using the Hartung-Knapp method,[18] which is more conservative and recommended for meta-analyses with few studies and low event rates. A sensitivity analysis leave-one-out was conducted for all outcomes to assess the influence of each individual study on the pooled estimates. For the primary outcome (function), an additional sensitivity analysis

<< SUMÁRIO

was conducted using the alternative Lysholm score reported by one of the studies. The funnel plot analysis was used to investigate potential publication bias in the primary outcome. Finally, meta-regression analyses were performed for the functional outcome to assess the influence of covariates such as the average age of participants, fracture severity (Meyers and McKeever classification), surgical access route (arthroscopic vs open), and sex distribution.

RESULTS

After the selection process, four articles were included in our study.[9,10,19,20] The flowchart detailing the selection process of the studies and the reasons for exclusion can be found in Figure 1.

Characteristics of the studies

Four retrospective cohort studies were included, conducted in South Korea[10], the United States[9], Europe (Finland, France, and Italy)[19], and Turkey,[20] with follow-up periods ranging from 1 to 2 years. In total, 186 patients were evaluated: 93 treated with screw fixation (SW) and 93 with suture fixation (SF). In the SW group, 63 patients were male (67.7%) and 30 were female (32.3%); in the SF group, 61 were male (65.6%) and 32 were female (34.4%). The average age was 14.5 ± 3.8 years in the SW group and 14.1 ± 4.1 years in the SF group. All fractures were classified according to the Meyers and McKeever system.[9,10,19,20] In the SW group, 1.1% were type I, 38.2% type II, and 60.1% type III. In the SF group, 29.0% were type II and 70.1% type III, with no cases of type I. The time between trauma and surgical intervention was reported in two studies (Seon et al., 2009[10] and Ercan et al., 2024[20]), with an average of 13.13 ± 2.99 days in the SW group and 20.52 ± 3.51 days in the SF group. The arthroscopic technique was predominant in both groups, being used in 97.2% of SW cases and 96% of SF cases.

Three studies reported the postoperative immobilization protocol, with full extension adopted for all patients.[10,19,20] The average duration of immobilization was 3.6 weeks, ranging from 2 to 4.5 weeks among the studies. More details of the included studies are presented in Table 1.

Meta-analysis of the included studies

Knee function

Three studies were included for this outcome, totaling 118 patients. Jääskelä et al. (2023)[19] and Ercan et al. (2024)[20] used the subjective IKDC scale, while Seon et al. (2009)[10] employed the Lysholm score to assess postoperative knee function. There was no statistically significant difference between the SW and SF groups (SMD = −0.07; 95% CI: −0.43 to 0.29; I2: −0.87 to 0.72; p = 0.702; I^2 = 0%) (Figure 2).

The leave-one-out sensitivity analysis demonstrated that the exclusion of any individual study did not substantially alter the overall estimate, which remained statistically non-significant (Figure S1). A sensitivity analysis was also conducted using the Lysholm score additionally reported by Ercan et al. (2024),[20] with equally consistent and non-significant results (SMD = 0.03; CI95%: −0.36 to 0.42; IQR: −0.99 to 1.04; p = 0.885; I^2 = 17.8%) (Figure S2). The funnel plot analysis revealed an asymmetric distribution of studies, suggesting possible publication bias (Figure S3).

The meta-regression did not identify statistically significant associations between functional outcomes and the evaluated covariates: (1) mean age of participants (Figure S4 and Table S1); (2) mean severity of fracture according to Meyers and McKeever (Figure S5 and Table S2); (3) proportion of patients treated via arthroscopic or open methods (Figure S6 and Table S3); (4) proportion of male patients (Figure S7 and Table S4).

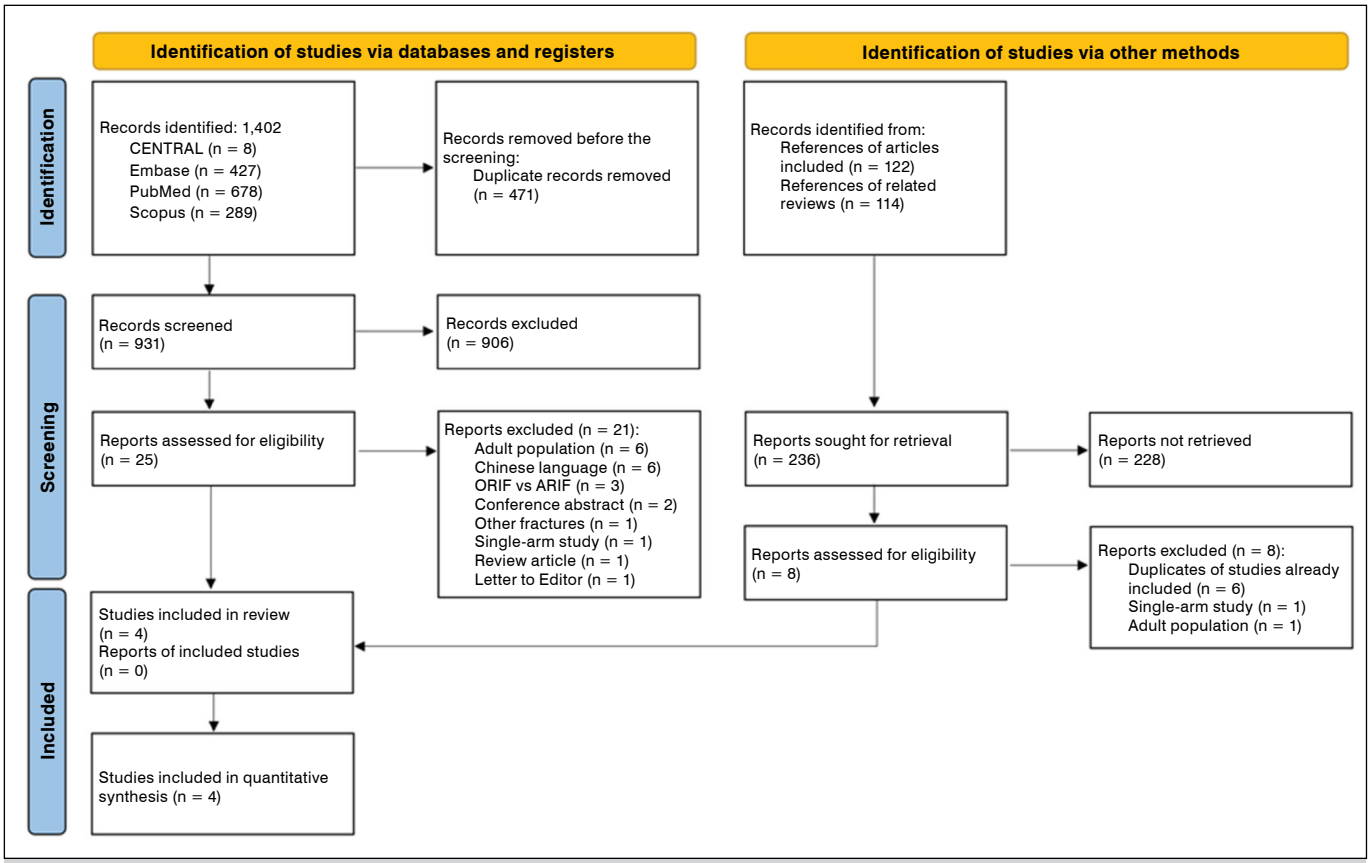


Figure 1. Flowchart of the included studies.

Table 1. Main characteristics of the included studies.

Study	Intervention		Design (Country)	Follow-up	Number of patients (M/F)		Age (years)		M&M Classification (I/ II/ III)		TIS (d)†		Surgical reduction approach (A/O)‡		Immobilization
	SW	SF			SW	SF	SW	SF	SW	SF	SW	SF	SW	SF	
Seon et al. (2009)	3.5–4.5 mm cannulated screw (1–2 units)	Non-absorbable suture	Retrospective cohort (South Korea)	2 years	10/6	13/04	17.7 ± 5.3	17.6 ± 7.2	0/6/10	0/5/12	29.8 ± 3.39	29.66 ± 3.34	32/ 3	24/4	Full extension with a long leg cast for 3 weeks
Callanan et al. (2019)	3.5–4.5 mm cannulated screw	Transosseous suture (Orthocord/ PDS/FiberWire No. 2/5)	Retrospective cohort (EUA)	1 year	27/08	22/11	11.2 ± 3.2	12.4 ± 2.5	0/9/22	0/5/28	ND	ND	31/0	33/0	ND
Jääskelä et al. (2023)	SW (screw type, diameter or implant not specified)	Suture technique/type not specified	Retrospective cohort (Finland, France, and Italy)	2 years	17/12	18/14	11.5 ± 2.3	10.9 ± 2.7	1/26/34	0/12/20	ND	ND	29/0	32/0	In full extension for 4.5 weeks
Ercan et al. (2024)	Headless compression screw (diameter/ material not specified)	Transosseous fixation with 2 Ultrabraid sutures	Retrospective cohort (Turkey)	2 years	9/4	8/3	10.8 ± 2.0	10.4 ± 2.1	0/5/8	0/5/6	6.0 ± 2.39	6.39 ± 3.76	13/0	12/0	In full extension with brace for 2 weeks.

Legends: SW = Screw fixation; SF = Suture fixation; M/F = Male/Female; M&M = Meyers and McKeever; I/II/III = Fracture classification I, II, III; TIS = Time from injury to surgery; d = days; A/O = Arthroscopic/*Open*; ND = Not described/Not defined; PDS = Polydioxanone; No. = Number. Mean ± Standard Deviation. Number of patients undergoing arthroscopy or open surgery.

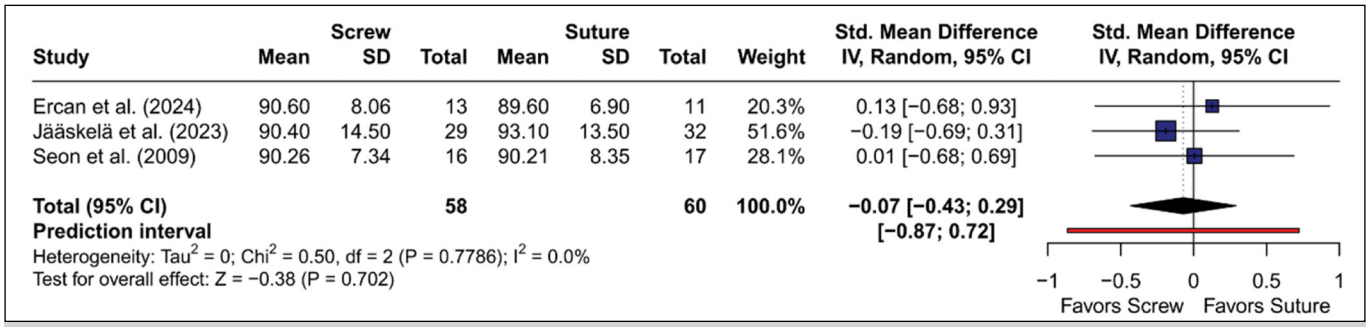


Figure 2. Forest plot of the meta-analysis for the functional outcome measured by the IKDC score.

Operative time

Three studies, totaling 124 patients,[9,10,20] were included for this outcome. The analysis, based on the average operative time in minutes, revealed a statistically significant difference in favor of the screw group (MD = –9.17; CI95%: –11.29 to –7.05; IQR: –13.82 to –4.52; $p < 0.001$; $I^2 = 66.1\%$) (Figure S8).

The leave-one-out sensitivity analysis showed that the exclusion of Seon et al. (2009)[10] resulted in an MD of –8.93 minutes (CI95%: –11.42 to –6.44; $I^2 = 72.2\%$), while the exclusion of Ercan et al. (2024)[18] generated an MD of –8.56 (CI95%: –12.70 to –4.42; $I^2 = 82.7\%$). On the other hand, the exclusion of Callanan et al. (2019)[] increased the effect size (MD = –29.45; CI95%: –45.97 to –12.93; $I^2 = 82.7\%$), while maintaining statistical significance in all scenarios (Figure S9).

Postoperative complications

Knee instability

Three studies, totaling 125 patients, assessed postoperative knee instability through the Lachman test. Em Seon et al. (2009)[10], patients with grade 1 or 2 were considered unstable, with no cases of higher grades. Ercan et al. (2024)[18] included only patients with grade 1 instability. Callanan et al. (2019)[9] described the cases as "unstable," without detailing the grades of the Lachman test. There was no statistically significant difference between the SW and SF groups (RR = 0.90; CI95%: 0.46 to 1.77; I^2: 0.21 to 3.95; $p = 0.770$; $I^2 = 0\%$) (Figure S10). The sensitivity analysis confirmed the stability of the estimate (Figure S11).

Failure to return to sport

Three studies, with 162 patients,[9,10,19] were included for this outcome. The analysis, based on the proportion of patients who did not return to sport, found no statistically significant difference between the groups (RR = 1.83; CI95%: 0.75 to 4.48; I^2: 0.26 to 13.06; $p = 0.190$; $I^2 = 0\%$) (Figure S12). The sensitivity analysis *leave-one-out* confirmed the robustness of the findings (Figure S13).

Reoperation

Four studies, totaling 186 patients, were included.[9,10,19,20] The analysis was based on the proportion of patients who required reoperation due to postoperative complications compared to those who did not require it. A statistically significant difference was observed between the groups, with a higher risk of reoperation in the SW group (RR = 1.78; 95% CI: 1.04 to 3.06; I^2: 0.86 to 3.68; $p = 0.040$; $I^2 = 0\%$) (Figure S14). The sensitivity analysis *leave-one-out* confirmed the robustness of the findings, as the exclusion of any individual study did not substantially alter the overall estimate (Figure S15).

<< SUMÁRIO

Implant removal

Three studies, with 125 patients, were included for review.[9,10,20] The analysis was based on the proportion of patients who underwent implant removal compared to those who did not in each group. A statistically significant difference was observed, with a higher risk associated with the SW group (RR = 6.25; 95% CI: 1.19 to 32.90; IP: 0.70 to 55.79; p = 0.040; I^2 = 0%) (Figure S16). The sensitivity analysis leave-one-out showed that the exclusion of Seon et al. (2009)[10] or Callanan et al. (2019)[9] did not substantially alter the overall estimate, which remained statistically significant. In contrast, the exclusion of Ercan et al. (2024)[18] resulted in an even greater risk of implant removal in the SW group (RR = 7.21; 95% CI: 1.80 to 28.87; I^2 = 0%) (Figure S17).

Assessment of the risk of bias of the included studies

The assessment of the risk of bias of the included studies, conducted with the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) tool, is summarized in Figures 3A and 3B.

Confounding bias

One study presented a moderate risk due to incomplete control of confounding factors[9]. Three studies were classified as having a severe risk: in Ercan et al. (2024)[20], the type of fixation was determined by the year of surgery, introducing temporal bias and confusion related to the evolution of the technique; in Jääskelä et al. (2023)[19], treatment varied by center and period, which may reflect institutional preferences and changes in practices; and in Seon et al. (2009)[10], the choice of fixation method changed over time, raising concerns about confusion arising from changes in clinical practice.

Selection of participants

Seon et al. (2009)[10] was considered to be of low risk. The other studies were classified as having a moderate risk. Callanan et al. (2019)[9] e Ercan et al. (2024)[20] excluded patients due to inadequate follow-up or associated injuries, which may have introduced selection bias. Jääskelä et al. (2023)[19] presented a low response rate (53%), which may affect representativeness.

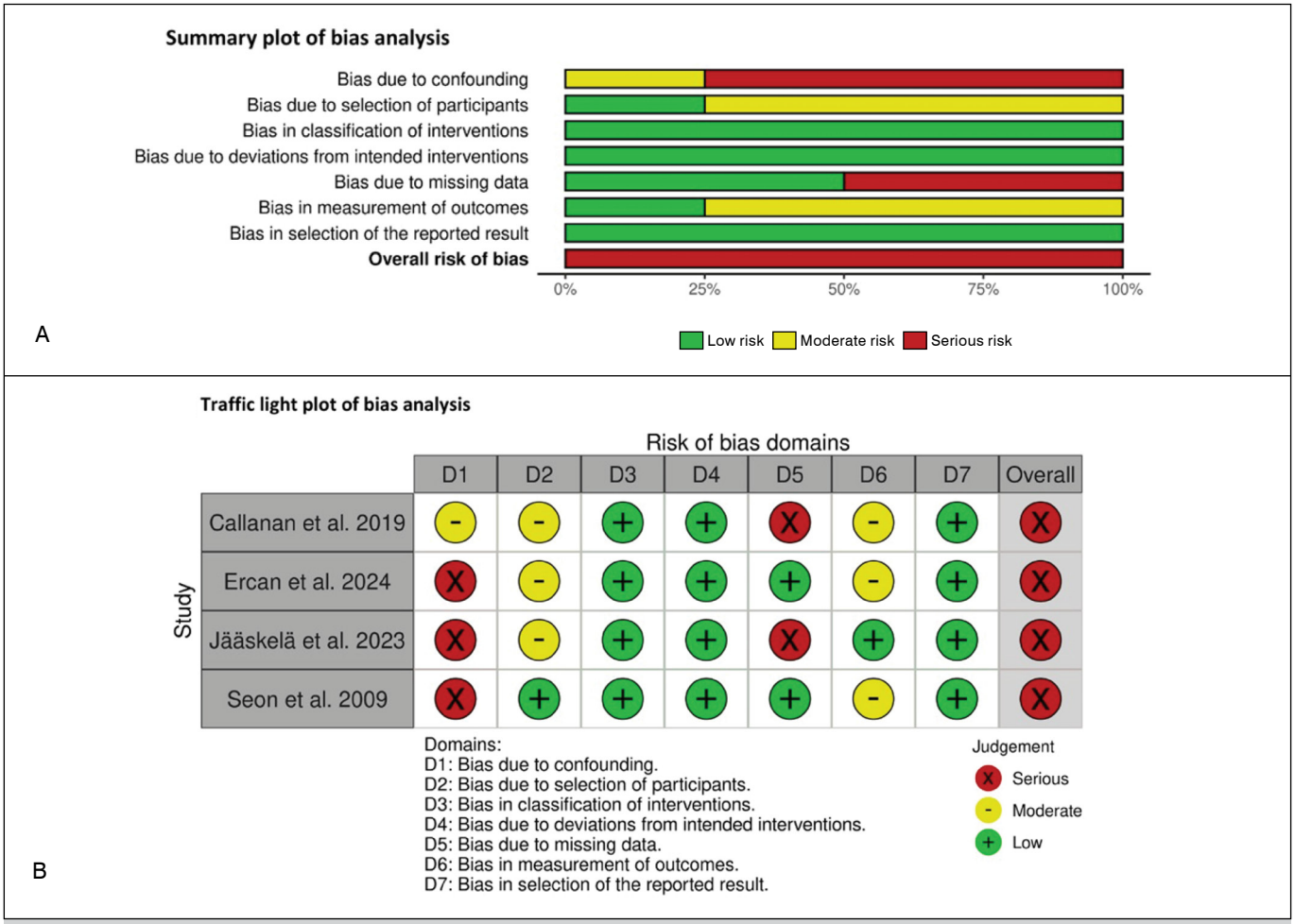


Figure 3. A and B. Summary and traffic graph of the bias analysis using the Risk of Bias in Non-Randomized Studies – of Interventions (ROBINS-I) tool.

Interventions

All studies were assessed as having a low risk of bias regarding the classification of interventions and deviations from intended interventions (items 3 and 4 of ROBINS-I)[9,10,19,20].

Missing data

Ercan et al. (2024)[20] e Seon et al. (2009)[10] were classified as low risk. Callanan et al. (2019)[9] e Jääskelä et al. (2023)[19] presented a serious risk due to high loss to follow-up, which may have influenced the results if the missing data were not random.

Outcomes

The article by Jääskelä et al. (2023)[19] was considered low risk. The other three studies presented a moderate risk due to the lack of blinding and the assessment of subjective outcomes by the surgeon responsible for the treatment.[9,10,20]

Selective reporting

All studies were assessed as having a low risk for selective reporting of outcomes, with consistent presentation of the pre-planned outcomes.[9,10,19,20]

DISCUSSION

This systematic review with meta-analysis demonstrated that, in pediatric patients with tibial spine fractures, both screw fixation and suture fixation result in comparable postoperative functional outcomes. However, screw fixation was associated with a shorter operative time, at the cost of a significantly higher risk of reoperation and implant removal. These findings proved robust in sensitivity analyses and were not significantly influenced by clinical variables such as age, sex, fracture severity, or surgical technique, as evidenced by meta-regression models. The funnel plot analysis for the primary functional outcome revealed an asymmetric distribution, with a notable absence of small studies favoring screw fixation. This pattern may suggest publication bias, particularly in contexts where neutral or favorable results for suture fixation are more frequently published. However, given the limited number of included studies, this evidence should be interpreted with caution as asymmetries in funnel plots with small samples may occur by chance. As highlighted in the Cochrane Handbook for Systematic Reviews of Interventions, funnel plots with fewer than 10 studies have limited reliability in detecting true publication bias[21]. Nevertheless, the wide prediction intervals observed suggest that the estimated effects may vary substantially in future studies, indicating that the differences between techniques may be modified as new evidence becomes available.

A previous systematic review supports the findings found in the present study. The review conducted by Chang et al. (2022)[22] included five observational studies with 184 patients—primarily adolescents and young adults—and identified a higher incidence of reoperations and implant removals in the screw fixation group, despite no significant differences in functional outcomes. Additionally, high rates of return to high-impact sports were reported with suture fixation, even in type II to IV fractures, suggesting satisfactory long-term functional recovery. Although derived from an older population, these findings present points of convergence with the present analysis. The higher rates of reoperation and adverse events reported by Chang et al. (2022)[22] may reflect anatomical, biomechanical, or scar response differences between older patients and children, reinforcing the need for specific studies for the pediatric population.

Gans et al. (2014)[23], in a systematic review focused exclusively on pediatric patients, analyzed 451 cases of tibial eminence fractures. The authors reported that type III and IV fractures were associated with greater loss of range of motion, as well as a higher complication rate with screw fixation. On the other hand, suture fixation showed high rates of return to sports but was associated with adverse events in up to 30% of cases. In our research, no association was observed between the type of fracture and functional outcomes in the meta-regressions. Still, the findings of Gans et al. (2014)[23] directly corroborate the results we found, regarding the functional equivalence between techniques and lower incidence of reoperations and complications with suture fixation. This consistency among studies strengthens the evidence that suture fixation presents a more favorable safety profile in this population, possibly due to the greater regenerative capacity and lower joint stiffness observed in children.

A randomized clinical trial with 90 young adults demonstrated superior functional performance with suture fixation, reflected in higher scores on the IKDC and Lysholm, along with the absence of reoperations in this group, in contrast to seven reinterventions in the screw group[24]. Similarly, Qu et al. (2022)[25] reported a lower rate of adverse events with suture fixation in a cohort of 69 patients, although functional outcomes were similar between the groups. These findings reinforce the safety profile of suture fixation and suggest that such benefits may extend to age groups beyond the pediatric.

Single-arm observational studies also provide relevant insights. Yuan et al. (2015)[26] and Çağlar et al. (2020)[27], both with pediatric cohorts treated with suture fixation, reported good functional outcomes, low complication rates, and satisfactory return to activities. Also, Lutz et al. (2021)[28], in a cohort of 23 patients, observed satisfactory ligament stability and return to high-impact sports, despite complications in 30% of cases. Regarding screw fixation, Shin et al. (2018)[29] reported implant-related complications in 27 pediatric patients, while Wiegand et al. (2014)[30] described preservation of short-term function in eight children treated with Herbert screws. Although these studies do not have comparative groups, their findings contribute to outlining the clinical profile and complications associated with each technique.

Limitations of the study

This study has limitations that must be considered. All included articles are of the retrospective cohort type, with inherent risk of selection bias, confounding, and lack of standardization in data collection. We used the ROBINS-I tool to assess the risk of bias and conducted meta-regression analyses to explore potential sources of heterogeneity; however, the existence of residual confounding cannot be completely ruled out.

The small number of studies and participants included may have limited statistical power, especially for secondary outcomes. We employed conservative statistical approaches, including continuity correction, the Hartung-Knapp method, and prediction intervals to mitigate this impact. The study by Seon et al. (2009)[10] included both adults and children, although the population was predominantly pediatric. The sensitivity analysis excluding this study did not show a significant change in the overall effect estimate. Furthermore, the meta-regression using the mean age of participants showed that age did not influence functional outcomes.

The absence of high-level randomized clinical trials limits the ability to establish causal relationships. However, given the scarcity of high-quality comparative data in this specific population, our synthesis of observational evidence provides relevant insights to support clinical practice. Finally, asymmetry was observed in the funnel plot for the functional outcome, suggesting possible publication bias. Due to the limited number of studies, the Egger test could not be applied; however, we employed funnel plots with significance contours to assist in interpretation.

Therefore, we adopted rigorous methodological practices, including prior registration of the protocol, independent and duplicate data extraction and evaluation, as well as multiple sensitivity analyses, which strengthens the reliability of our findings.

CONCLUSION

This systematic review with meta-analysis demonstrated that suture fixation was associated with lower rates of reoperation and complications compared to screw fixation, while maintaining similar functional outcomes in the treatment of tibial spine fractures. These findings were consistent in sensitivity analyses and were not influenced by the clinical variables assessed through meta-regression. However, the prediction intervals suggest that these effects may vary with the publication of new studies, highlighting the presence of residual uncertainty. Although the available data indicate a potentially more favorable safety profile for suture fixation in pediatric patients, this evidence was based on observational studies. Better randomized clinical trials are needed to confirm these findings and guide the choice of the most appropriate surgical technique for tibial spine fractures in children.

CONTRIBUTIONS OF THE AUTHORS

Each author made a personal and significant contribution to the development of this article. Conceptualization: PEDL, JGS; AAN. Data Protection: AAN, MVRF, VMR. Methodology: PEDL, ETD, AAN. Writing original draft: PEDL, JGS, VMR. Supervision: ETD, MVRF. Writing – review and editing: PEDL, JGS, ETD.

DATA AVAILABILITY DECLARATION

The materials underlying this research are publicly available on the Figshare platform, through the following DOIs: https://doi.org/10.6084/m9.figshare.31915674 referring to Supplemental Material 2, containing the complementary figures and tables; and https://doi.org/10.6084/m9.figshare.31915419, referring to Supplemental Material 1, containing the complete search strategy used in the review.

REFERENCES

1. Meyers MH, Mckeever FM. Fracture of the intercondylar eminence of the tibia. J Bone Joint Surg Am. 1970;52(8):1677-84.
2. LaFrance RM, Giordano B, Goldblatt J, Voloshin I, Maloney M. Pediatric Tibial Eminence Fractures: Evaluation and Management. J Am Acad Orthop Surg. 2010;18(7):395-405. doi: 10.5435/00124635-201007000-00002.
3. May JH, Levy BA, Guse D, Shah J, Stuart MJ, Dahm DL. ACL Tibial Spine Avulsion: Mid-Term Outcomes and Rehabilitation. Orthopedics. 2011;34(2):89. doi: 10.3928/01477447-20101221-10.
4. Ahn JH, Yoo JC. Clinical outcome of arthroscopic reduction and suture for displaced acute and chronic tibial spine fractures. Knee Surg Sports Traumatol Arthrosc. 2005;13(2):116-21. doi: 10.1007/s00167-004-0540-6.
5. Hunter RE, Willis JA. Arthroscopic fixation of avulsion fractures of the tibial eminence: technique and outcome. Arthroscopy. 2004;20(2):113-21. doi: 10.1016/j.arthro.2003.11.028.
6. Bogunovic L, Tarabichi M, Harris D, Wright R. Treatment of Tibial Eminence Fractures: A Systematic Review. J Knee Surg. 2014;28(3):255-62.
7. Johnstone TM, Baird DW, Cuellar-Montes A, van Deursen WH, Tompkins M, Ganley TJ, et al. Screws or Sutures? A Pediatric Cadaveric Study of Tibial Spine Fracture Repairs. Am J Sports Med. 2023;51(10):2589-95. doi: 10.1177/03635465231181059.
8. Mahar AT, Duncan D, Oka R, Lowry A, Gillingham B, Chambers H. Biomechanical Comparison of Four Different Fixation Techniques for Pediatric Tibial Eminence Avulsion Fractures. J Pediatr Orthop. 2008;28(2):159-62. doi: 10.1097/BPO.0b013e318164ee43.
9. Callanan M, Allen J, Flutie B, Tepolt F, Miller PE, Kramer D, et al. Suture Versus Screw Fixation of Tibial Spine Fractures in Children and Adolescents: A Comparative Study. Orthop J Sports Med. 2019;7(11):2325967119881961. doi: 10.1177/2325967119881961.
10. Seon JK, Park SJ, Lee KB, Gadikota LR, Kozanek M, Oh LS, et al. A clinical comparison of screw and suture fixation of anterior cruciate ligament tibial avulsion fractures. Am J Sports Med. 2009;37(12):2334-9. doi: 10.1177/0363546509341031.
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow C, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372:71-80. doi: 10.1136/bmj.n71.
12. Lavigne PED, Filho MVR, Ribeiro, VM, de Souza JG, Nascimento AA, Dobashi ET. SUPPLEMENTARY MATERIAL 2 – FIGURES AND TABLES ("Screw versus Suture in the Fixation of Tibial Spine Fractures: A Systematic Review, Meta-Analysis, and Meta-Regression"). Figshare. 2026. Online resource. doi: 10.6084/m9.figshare.31915674.v1.
13. Lavigne PED, Filho MVR, Ribeiro, VM, de Souza JG, Nascimento AA, Dobashi ET. SUPPLEMENTARY MATERIAL 1 – SEARCH STRATEGY ("Screw versus Suture in the Fixation of Tibial Spine Fractures: A Systematic Review, Meta-Analysis, and Meta-Regression"). Figshare. 2026. Online resource. doi: 10.6084/m9.figshare.31915419.v3.
14. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210. doi: 10.1186/s13643-016-0384-4.
15. Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. New York: ACM; 2014. p. 1-10.
16. Sterne JA, Hernán MA, Reeves BC, Savovi□ J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi: 10.1136/bmj.i4919.
17. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of Two Methods to Detect Publication Bias in Meta-analysis. JAMA. 2006;295(6):676-80. doi: 10.1001/jama.295.6.676.
18. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014;14:25. doi: 10.1186/1471-2288-14-25.
19. Jääskelä M, Turati M, Lempainen L, Bremond N, Courvoisier A, Henri A, et al. Long-term Outcomes of Tibial Spine Avulsion Fractures After Open Reduction With Osteosuturing Versus Arthroscopic Screw Fixation: A Multicenter Comparative Study. Orthop J Sports Med. 2023;11(6):23259671231176991. doi: 10.1177/23259671231176991.
20. Ercan N, Arıcan G, □ibar K, Özmeriç A, □ltar S. Clinical and Functional Outcomes of Suture Versus Headless Screw Fixation for Tibial Eminence Fractures in Children. Am J Sports Med. 2024;52(4):948-55. doi: 10.1177/03635465241227440.
21. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). Cochrane, 2024. Available from www.cochrane.org/handbook.
22. Chang C-J, Huang T-C, Hoshino Y, Wang CH, Kuan FC, Su WR, et al. Functional Outcomes and Subsequent Surgical Procedures After Arthroscopic Suture Versus Screw Fixation for ACL Tibial Avulsion Fractures: A Systematic Review and Meta-analysis. Orthop J Sports Med. 2022;10(4):23259671221085945. doi: 10.1177/23259671221085945.
23. Gans I, Baldwin KD, Ganley TJ. Treatment and Management Outcomes of Tibial Eminence Fractures in Pediatric Patients. Am J Sports Med. 2014;42(7):1743-50. doi: 10.1177/0363546513508538.
24. Jain S, Modi P, Dayma RL, Mishra S. Clinical outcome of arthroscopic suture versus screw fixation in tibial avulsion of the anterior cruciate ligament in skeletally mature patients. J Orthop. 2023;35:7-12. doi: 10.1016/j.jor.2022.10.006.
25. Qu H, Meng Q, Sun Q, Du D, Zhang Q. Arthroscopic fixation for tibial eminence fractures: A clinical retrospective study of cannulated screws versus transosseous anchor knot fixation techniques with suture anchors. Knee. 2022;35:105-13. doi: 10.1016/j.knee.2022.02.002.
26. Yuan Y, Huang X, Zhang Y, Wang Z. Treatment of tibial eminence fractures with arthroscopic suture fixation technique: a retrospective study. Int J Clin Exp Med. 2015;15(8)8:13797-803.
27. Ça□lar C, Ya□ar H, Emre F, U□urlu M. Mid-term Outcomes of Arthroscopic Suture Fixation Technique in Tibial Spine Fractures in the Pediatric Population. Ulus Travma Acil Cerrahi Derg. 2021;27(5):571-6. doi: 10.14744/tjtes.2020.25905.
28. Lutz PM, Geyer S, Winkler PW, Irger M, Berthold DP, Feucht MJ, et al. Reliable ligamentous stability and high return-to-sport rates after arthroscopic reduction and internal fixation of tibial eminence fractures. Arch Orthop Trauma Surg. 2021;142(12):3623-31. doi: 10.1007/s00402-021-03961-6.
29. Shin CH, Lee DJ, Choi IH, Cho TJ, Yoo WJ. Clinical and radiological outcomes of arthroscopically assisted cannulated screw fixation for tibial eminence fracture in children and adolescents. BMC Musculoskelet Disord. 2018;19(1):41. doi: 10.1186/s12891-018-1960-7.
30. Wiegand N, Naumov I, Vámhidy L, Nöt LG. Arthroscopic treatment of tibial spine fracture in children with a cannulated Herbert screw. Knee. 2014;21(2):481-5. doi: 10.1016/j.knee.2013.12.004.

DOI: http://dx.doi.org/10.1590/1413-785220263404e300345

USE OF ARTIFICIAL INTELLIGENCE FOR THE EVALUATION OF THE ENTRANCE EXAMINATION OF THE BRAZILIAN SOCIETY OF SHOULDER AND ELBOW SURGERY

USO DA INTELIGÊNCIA ARTIFICIAL PARA AVALIAÇÃO DO EXAME DE INGRESSO DA SOCIEDADE BRASILEIRA DE CIRURGIA DO OMBRO E COTOVELO

Caio Henrique Kenchian[1], *Karolina Stephany Pereira Ferreira*[1], *Lucas Pereira Sarmento*[1], *Marcelo Silveira Teixeira Filho*[1], *Renato Aroca Zan*[1], *Marcel Jun Sugawara Tamaoki*[1]

1. Universidade Federal de Sao Paulo, Escola Paulista de Medicina, Departamento de Ortopedia e Traumatologia, Sao Paulo, SP, Brazil.

ABSTRACT

Introduction: AI is increasingly used for medical education and assessment, yet effectiveness on specialized exams remains underexplored. We aim to compare multiple AI models with national human averages on SBCOC entrance exams (2021–2023), evaluate answer accuracy and citation reliability, and contrast models. Methods: Five models—ChatGPT-4 (standard and literature-trained), ChatGPT-o1-pro, Gemini, and Meta Llama 3.1—answered official SBCOC exams (50 items/year) individually using a unified prompt and identical images. Each response included one choice (A–D) and a cited source. Outcomes were accuracy and source reliability (peer-reviewed articles/textbooks versus websites). Statistics used chi-square tests and one-way ANOVA with Tukey post-hoc ($p<0.05$). Results: Across 150 questions, ChatGPT-o1-pro led (66%, 62%, 68%) and exceeded human national averages each year ($p<0.05$). Gemini (40%, 28%, 38%) and Meta Llama 3.1 (32%, 44%, 50%) underperformed relative to humans, while both ChatGPT-4 versions hovered near the ≥50% pass threshold without significant differences. Regarding sources, ChatGPT-o1-pro and Llama predominantly cited articles or books, ChatGPT-4 alternated between literature and websites, and Gemini did not specify references. Conclusions: ChatGPT-o1-pro outperformed the national human average and mostly used credible sources; ChatGPT-4 matched humans, while Gemini and Meta Llama 3.1 lagged. ***Level of evidence IV; Case series.***

Keywords: Educational Measurement; Education, Medical; Artificial Intelligence; Orthopedics; Natural Language Processing.

RESUMO

Introdução: A IA é cada vez mais usada na educação e avaliação médicas, mas sua efetividade em exames especializados permanece pouco explorada. O objetivo foi comparar modelos de IA às médias humanas nos exames de ingresso da SBCOC (2021–2023), avaliar acurácia e confiabilidade das citações e contrastar modelos. Métodos: Cinco modelos—ChatGPT-4 (padrão e treinado com literatura), ChatGPT-o1-pro, Gemini e Meta Llama 3.1—responderam às provas da SBCOC (50 itens/ano) individualmente, com prompt unificado e imagens idênticas. Cada resposta continha uma alternativa (A–D) e uma fonte. Desfechos: acurácia e confiabilidade das fontes (artigos/livros vs. sites). Estatística: qui-quadrado e ANOVA de uma via com pós-teste de Tukey ($p<0,05$). Resultados: Em 150 questões, o ChatGPT-o1-pro liderou (66%, 62%, 68%) e superou médias humanas em todos os anos ($p<0,05$). Gemini (40%, 28%, 38%) e Meta Llama 3.1 (32%, 44%, 50%) abaixo dos humanos; ambas as versões do ChatGPT-4 próximas ao limiar de aprovação ≥50%, sem diferenças significativas. Quanto às fontes, ChatGPT-o1-pro e Llama citaram predominantemente artigos/livros; o ChatGPT-4 alternou entre literatura e sites; Gemini não especificou referências. Conclusões: O ChatGPT-o1-pro superou a média humana e usou majoritariamente fontes confiáveis; o ChatGPT-4 igualou os humanos; Gemini e Meta Llama 3.1 ficaram aquém. ***Nível de evidência IV; série de casos.***

Descritores: *Avaliação Educacional; Educação Médica; Inteligência Artificial; Ortopedia; Processamento de Linguagem Natural.*

Citation: Kenchian CH, Ferreira KSP, Sarmento LP, Teixeira Filho MS, Zan RA, Tamaoki MJS. Use of artificial intelligence for the evaluation of the entrance examination of the brazilian society of shoulder and elbow surgery. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 5. Available from URL: http://www.scielo.br/aob.

All authors declare no potential conflict of interest related to this article.

The study was conducted at Universidade Federal de Sao Paulo, Escola Paulista de Medicina, Department of Orthopedics and Traumatology, Sao Paulo, SP, Brazil.
Correspondence: Caio Henrique Kenchian. 140, Rua Benta Pereira, ap 182, Sao Paulo, SP, Brazil, 02451000. ckenchian@gmail.com

Article received on 09/23/2025 approved on 03/10/2026.
Handling Editor: Américo Zoppi Filho

<< SUMÁRIO

INTRODUCTION

Artificial intelligence (AI) has undergone remarkable evolution since its early applications in the 1950s, when it was limited to logic-based and symbolic reasoning systems, to the complex machine learning models we know today. In the field of medicine, AI began to gain relevance with the development of specialized systems, such as MYCIN in the 1970s, one of the first to assist in the diagnosis of bacterial infections[1]. However, it was never used clinically due to technical and trust limitations[2]. Since then, AI techniques have evolved considerably, driven by improvements in processing power and the emergence of deep neural networks, enabling AI to perform tasks that once seemed unattainable.

Today, AI is successfully applied in various medical specialties[3-5]. However, not all attempts to integrate AI into medicine have been successful. Many systems struggle to handle the complexity and variability of real-world clinical conditions[2]. Additionally, the lack of high-quality data and the resistance of healthcare professionals to fully adopt these technologies remain recurring barriers.

In contrast, in the field of medical education, AI adoption has been more widely accepted[6]. Tools like ChatGPT provide support to students, especially in simulation scenarios and problem-based tutorials, offering a more immersive and personalized learning experience[7]. However, much of the information provided by these tools can be inaccurate or sometimes inappropriate. Therefore, these advancements need to be continuously evaluated to ensure that AI is used safely and effectively in medical practice.

Natural language processing (NLP) models, particularly large language models (LLMs) like ChatGPT, have made significant advancements in recent years. Various studies[8-10] have already explored these models, demonstrating not only their high capability in solving medical exams but also highlighting the superiority of newer models.

In addition to ChatGPT, other NLP models have also gained relevance in medicine. Gemini, developed by Google DeepMind, combines language processing with the analysis of large volumes of clinical data to provide more accurate diagnoses and personalized treatments[11]. It is important to note that this study employed the standard publicly available Gemini model, not the medically specialized 'Med-Gemini' described by Saab et al[11]. These models stand out due to their unique architectures and learning capabilities; however, it is essential to continuously monitor their accuracy and safety to ensure their ethical and secure use.

The use of AI in medical examinations also raises concerns about its reliability and accuracy. Although natural language models have shown promising performance, a systematic evaluation is necessary to ensure that the results are consistent and accurate[12]. Studies validating the effectiveness of these models in clinical simulations and across different medical specialties are essential to ensure that these technologies effectively contribute to the training of future specialists[13].

Therefore, although the advancements are remarkable, there are still important gaps to be addressed. More detailed studies are needed to assess AI performance in image interpretation and ensure the accuracy and robustness of the information provided. These are areas that require improvement for AI to be safely and effectively integrated into both medical education and clinical practice.

The justification for this study lies in the increasing use of artificial intelligence (AI) as a support tool in medical education and clinical practice. Although AI models such as ChatGPT, Gemini, and Meta Llama demonstrate great potential in solving complex problems, there is a gap in evaluating their performance compared to human candidates, especially in more specialized exams like those of the Brazilian Society of Shoulder and Elbow Surgery (SBCOC).

Thus, this study aims to compare the performance of AI models among themselves and in comparison to the national averages of human candidates, both in terms of response accuracy and the reliability of the sources used, in the entrance exams of the Brazilian Society of Shoulder and Elbow Surgery (SBCOC) from 2021 to 2023. It will also compare trained versions, with access to medical literature used as a reference, to untrained versions, with free access to the internet. Additionally, the evaluation of the AIs will include questions that involve image interpretation for resolution.

METHODS

Study Design and Overview

The evaluation of different Artificial Intelligence (AI) models was conducted using the official examinations of the Brazilian Society of Shoulder and Elbow Surgery (SBCOC) from 2021, 2022, and 2023. These admission examinations comprise fifty multiple-choice questions with four alternatives, only one of which is correct, and the SBCOC supplied the official exams, corresponding answer keys, and anonymized statistical data on candidate performance. Five AI models were evaluated—ChatGPT-4 in two modalities (a standard version with open internet access and an isolated version restricted exclusively to the exam's reference literature), ChatGPT-o1-pro, Gemini (public version, without the medical enhancements of Med-Gemini), and Meta Llama 3.1—enabling direct comparison between models with unrestricted access to external information and those limited to a closed reference set.

To standardize administration, a single prompt was used across all models: "I will provide a multiple-choice question on the topic of shoulder and elbow surgery, taken from an examination for orthopedic surgeons. I want you to select the best alternative to answer the question, using the information provided in the question as well as data from your database or the internet. I also want you to provide the source used to solve the question. The answer must be in the format of one alternative: A, B, C, or D".

Each question was presented individually, because administering the entire exam at once caused oversimplification and higher error rates; when questions included images, the same images were attached within each AI's interface to ensure identical visual content.

Model responses were assessed on two criteria: (1) accuracy, via direct comparison with the official SBCOC answer key; and (2) source usage, analyzing the reliability of cited references, with textbooks and peer-reviewed articles deemed reliable and websites or non-validated materials deemed unreliable.

Tests were conducted under standardized conditions: all models received identical instructions and were evaluated on their native user platforms within a short time frame to minimize performance variations from back-end updates, thus promoting consistency and comparability across models.

The collected data underwent statistical analysis with the objectives of comparing performance among AI models and against human performance (national average of SBCOC candidates across the three years) and assessing the reliability of cited sources, thereby ensuring methodological rigor for both answer-accuracy comparisons and critical appraisal of reference quality.

Statistical Analysis

To evaluate significant differences in AI performance among models and compared to human candidates across the study period (2021–2023), chi-square tests (χ^2) were applied to categorical variables (correct vs. incorrect answers), with a significance level set at $p<0.05$. Additionally, an analysis of variance (ANOVA) followed by Tukey's multiple comparison tests was used to compare average performance scores among the AI models. The tested

hypothesis was that significant differences would exist among the models and between the different exam years. All statistical analyses were conducted using JASP software, version 0.19.2 (2024), with a significance threshold of 0.05.

Ethical Considerations

This study was conducted in accordance with the ethical principles. It was approved by the Research Ethics Committee of the Federal University of São Paulo – Escola Paulista de Medicina (UNIFESP - EPM) and registered on the "Plataforma Brasil" under the CAAE number 81368024.9.0000.5505. The research involved the analysis of data from candidates who took the entrance exam of the Brazilian Society of Shoulder and Elbow Surgery. All data were anonymized prior to analysis to ensure the confidentiality and privacy of the participants. No identifiable personal information was used at any stage of the study.

RESULTS

The analyses were conducted under two circumstances. The first analysis employed the chi-square (χ^2) contingency test, which evaluates deviations in proportions and assesses differences in the observed proportions of correct answers among AI models compared to the established pass threshold of 50% (Table 1). The chi-square test identified significant differences among the AI models across the analyzed years (χ^2; $p<0.05$). In 2021, ChatGPT-o1-pro had the highest accuracy rate (66%), followed by ChatGPT-4 (60%), while Gemini (40%) and Meta Llama 3.1 (32%) performed worse and did not reach the passing threshold. In 2022, ChatGPT-o1-pro maintained the best performance (62%), while ChatGPT-4 (50%) and Meta Llama 3.1 (44%) had intermediate performance, and Gemini had the lowest score (28%). In 2023, ChatGPT-o1-pro again achieved the highest accuracy rate (68%), being the only AI to consistently outperform the others over the years.

The second analysis, focusing on individual model performances, was conducted using a one-way ANOVA without repetition, performed after a positive parametricity test, to verify whether there were differences in the geometric mean accuracy rates among the different AI models, including accuracy in text-based questions and image-based questions (Table 2). The geometric mean was selected due to the percentage-based nature of the data, as it is particularly suitable for averaging proportional rates or indices and mitigates the disproportionate impact of extreme values. Table 2 shows that ChatGPT-o1-pro achieved the highest average performance (65.3%), with a significant difference confirmed by ANOVA ($p=0.03$). In image-based questions, all AIs except Gemini had an average performance above 50%, with ChatGPT-o1-pro standing out (73.7%), being the only model to consistently exceed 50% accuracy in each individual exam (Table 3).

An inference was also made comparing candidates versus AIs, based on the percentage difference in accuracy rates, analyzed using the chi-square contingency test (Table 4). It is evident that ChatGPT-o1-pro consistently outperformed the national average of candidates in all years, being 14% superior in 2021 ($p=0.04$), 20% superior in 2022 ($p=0.03$), and 36% superior in 2023 ($p=0.029$). Meanwhile, Gemini and Meta Llama 3.1 were outperformed by the candidates in all years. ChatGPT-4, in both versions (trained and untrained), demonstrated statistically similar performance to the candidates.

Finally, as contributory elements for confirming the hypotheses of this study, an analysis was conducted to compare the bibliographic sources used by the AIs and their performance (Table 5). It was observed that ChatGPT-o1-pro and Meta Llama 3.1 exclusively used books and scientific articles across all three years. ChatGPT-4, on the other hand, alternated between websites and scientific literature, with an increase in the use of articles in 2023 (58%). The Gemini model did not specify the sources used, merely stating that it relied on articles, websites, and the internet in general, which were considered unreliable sources. The chi-square test (χ^2) indicated that the preference for sources varied significantly among the AIs.

Table 1. Percentage of correct answers among five different artificial intelligence models.

Year	AI Model	Correct Answers*	Correct	Passed	χ^2 (p-value)
2021	GPT-4	30	60%	Yes	>11.07 (p=0.03)
	GPT-4 (trained)	26	52%	Yes	
	GPT-o1-pro	33	66%	Yes	
	Gemini	20	40%	No	
	Meta Llama 3.1	16	32%	No	
2022	GPT-4	25	50%	Yes	>9.48 (p=0.01)
	GPT-4 (trained)	28	56%	Yes	
	GPT-o1-pro	31	62%	Yes	
	Gemini	14	28%	No	
	Meta Llama 3.1	22	44%	No	
2023	GPT-4	24	48%	No	>7.79 (p=0.01)
	GPT-4 (trained)	26	52%	Yes	
	GPT-o1-pro	34	68%	Yes	
	Gemini	19	38%	No	
	Meta Llama 3.1	25	50%	Yes	

* The calculated χ^2 value (df = 5) exceeded the critical value throughout the period, indicating a significant difference at the 5% level regarding the proportions of correct answers. Therefore, the null hypothesis (H_0), which suggested no differentiation between the AIs, is rejected.

DISCUSSION

Principal Results

This study evaluated AI performance on SBCOC medical exams (2021–2023) against the national average of human candidates, assessing answer accuracy and the use of reliable bibliographic

Table 2. Analysis of variance between the average correct answers for text questions between 2021 and 2023.

AI Model	Average Correct Answers (2021–2023) and 95% Confidence Interval*	ANOVA**
GPT-4	26.2 (24.2 - 27.8)	F > 6553 (p=0.03)
GPT-4 (trained)	26.6 (27.2 - 28.0)	
GPT-o1-pro	32.6 (29.8 - 37.3)	
Gemini	17.0 (15.2 - 18.0)	
Meta Llama 3.1	21.2 (19.2 - 23.2)	

* The confidence intervals for percentages were calculated using the standard error (SE) corrected by stratification: $\sqrt{p(1-p)/n}$, and with 95% confidence = $p' \pm 1.96 \times SE$. ** The ANOVA test revealed a significant difference between the AIs, confirmed by the Tukey post hoc test, which showed differences in the means, particularly with the GPT-o1-pro.

Table 3. Analysis of variance between the average correct answers for image questions between 2021 and 2023.

AI Model	Average* Correct Answers (2021–2023) and 95% Confidence Interval**	ANOVA***
GPT-4	2.0 (0.8 - 2.8)	F > 6553 (p=0.03)
GPT-4 (trained)	2.0 (1.0 - 3.0)	
GPT-o1-pro	2.6 (2.1 - 3.1)	
Gemini	1.0 (0.8 - 1.8)	
Meta Llama 3.1	2.0 (1.2 - 2.8)	

* Average of 3.66 image-based questions per exam/year. ** The confidence intervals for percentages were calculated using the standard error (SE) corrected by stratification: $\sqrt{p(1-p)/n}$, and with 95% confidence = $p' \pm 1.96 \times SE$. *** The ANOVA test revealed a significant difference between the AIs, confirmed by the Tukey post hoc test, which showed differences in the means, particularly with the GPT-o1-pro.

Table 4. Percentage difference between candidates' and AI's correct answers from 2021 to 2023.

Year	AI Model	Correct Answers	χ^2 (p-value)
2021	GPT-4	30	Performance 3.3% better for AI** (p=0.07)
	GPT-4 (trained)	26	Performance 11.5% better for candidates (p=0.06)
	GPT-o1-pro*	33	Performance 14% better for AI** (p=0.04)
	Gemini	20	Performance 45% better for candidates (p<0.01)
	Meta Llama 3.1	16	Performance 82% better for candidates (p<0.01)
	Students	29	-
2022	GPT-4	25	Performance 4% better for candidates (p=0.07)
	GPT-4 (trained)	28	Performance 7.2% better for AI** (p=0.07)
	GPT-o1-pro*	31	Performance 20% better for AI** (p=0.03)
	Gemini	14	Performance 86% better for candidates (p<0.01)
	Meta Llama 3.1	22	Performance 18% better for candidates (p<0.01)
	Students	26	-
2023	GPT-4	24	Performance 4.7% better for candidates (p=0.07)
	GPT-4 (trained)	26	Performance 4% better for AI** (p=0.06)
	GPT-o1-pro*	34	Performance 36% better for AI** (p=0.03)
	Gemini	19	Performance 31.7% better for candidates (p=0.02)
	Meta Llama 3.1	25	Similar performance
	Students	25	-

* The average number of correct answers for the AI GPT-o1-pro over the three years was 32.6 (65.2%), showing a significant difference ($p<0.05$, as assessed by χ^2) compared to the students› average of 26.6 (53%). ** The χ^2 test did not reveal a significant difference between the AIs and candidates, despite GPT-4 and GPT-4 Trained having better percentage performances. Candidates outperformed only with Gemini and Meta Llama 3.1 ($p<0.01$ in comparison to both, as assessed by χ^2).

Table 5. Percentage difference in the bibliographic sources used by different AIs to solve the questions.

Year	AI Model	Books and Articles	Websites	χ^2 (p-value) *
2021	GPT-4	23 (46%)	27 (54%)	p=0.03 (websites)
	GPT-o1-pro	50 (100%)	-	-
	Meta Llama 3.1	50 (100%)	-	-
2022	GPT-4	14 (28%)	36 (72%)	p=0.02 (websites)
	GPT-o1-pro	50 (100%)	-	-
	Meta Llama 3.1	50 (100%)	-	-
2023	GPT-4	29 (58%)	21 (42%)	p=0.04 (articles)
	GPT-o1-pro	50 (100%)	-	-
	Meta Llama 3.1	50 (100%)	-	-

* The calculated χ^2 value (df = 2) exceeded the tabled value.

sources. Regarding the passing threshold (≥50% correct answers), both trained ChatGPT-4 and ChatGPT-o1-pro met the criterion across all years; between them, ChatGPT-o1-pro outperformed the candidates and achieved the highest average number of correct answers, including on image-based questions, standing out particularly in 2022 and 2023 when candidate performance declined markedly. A model-specific analysis revealed a performance gap: Gemini and Meta Llama 3.1 obtained the lowest results relative to other AIs and to humans, while trained and standard ChatGPT-4 showed no significant difference from each other and did not differ significantly from the candidates.

Source behavior further differentiated models: ChatGPT-4 frequently relied on websites (a limitation due to lower reliability), ChatGPT-o1-pro consistently relied on trustworthy sources, and Gemini did not specify sources—citing only "articles, websites, and the internet in general"—hindering quality appraisal. Notably, although Meta Llama 3.1 performed worse than other AIs, it referred exclusively to ostensibly reliable sources; however, its author–book pairings were often inaccurate, indicating fabricated attributions and underscoring the need to carefully analyze AI-generated responses for invented information and references.

Study Limitations

Although the testing was conducted within a close time frame and on each model's standard usage platform to minimize system-related variability, the possibility of performance fluctuations due to back-end factors cannot be fully excluded. Updates to infrastructure, adjustments in model parameters, or system-level optimizations may have occurred during the evaluation period, introducing uncontrolled variation into the results. While efforts were made to standardize the conditions of application, these back-end variations remain an inherent limitation in the assessment of proprietary AI models, as researchers have no direct control or visibility over the dynamic changes that may take place in their underlying architectures.

Comparison with Prior Work

This review situates the present study within a growing literature showing strong AI performance on medical multiple-choice examinations and clarifies key limitations and contributions.

Prior work on the United States Medical Licensing Examination (USMLE)[10] reported that ChatGPT correctly answered more than 60% of questions, reaching the passing threshold for Steps 1 and 2 and suggesting knowledge comparable to a third-year medical student.

On the Orthopaedic In-Training Examination (OITE)[8], ChatGPT-3.5 answered 196 of 360 questions (54.3%), whereas ChatGPT-4 achieved 265 of 360 (73.6%)—equivalent to the performance of a fifth-year resident (PGY-5) and above the minimum passing score for Part I of the American Board of Orthopaedic Surgery Exam; notably, ChatGPT-4 provided verifiable sources in 87.9% of responses, versus 47.2% for the earlier version.

ChatGPT-4 also performed strongly on the Neurological Surgeons Self-Assessment Neurosurgery Exam (SANS)[9], attaining high accuracy across question categories (tumors, cerebrovascular conditions, trauma, pediatrics) and outperforming medical students, neurosurgery residents, and the national average of SANS users, implying it would likely pass the American Board of Neurological Surgery primary examination.

Additionally, Costa et al. reported that ChatGPT-4 answered 61.05% of questions correctly on the 2022 TEOT first-phase exam—surpassing the ≥60% passing threshold—reinforcing that LLMs can meet Brazilian specialist-exam benchmarks in orthopedics[14].

This article makes a distinct contribution by benchmarking multiple AI models (ChatGPT-4, ChatGPT-o1-pro, Gemini, Meta Llama 3.1, and a trained ChatGPT-4) against national human averages on SBCOC entrance exams (2021–2023), assessing not only overall accuracy but also the reliability of cited sources within a highly specialized surgical domain. The findings align with and extend prior literature by demonstrating model behavior in a narrower,

<< SUMÁRIO

domain-specific context, reinforcing the potential of advanced systems—particularly ChatGPT-o1-pro—as supplementary educational tools that aid preparation for specialized examinations and enhance learning outcomes, while underscoring the need for ongoing critical monitoring of AI-generated content and sources and for continuous refinement to ensure safe, effective use in educational and clinical settings.

CONCLUSION

The average performance of ChatGPT-o1-pro consistently exceeded the national average of human candidates, and it predominantly cited reliable sources; by contrast, ChatGPT-4 (across all versions) demonstrated average performance comparable to that of humans and consistent across versions, whereas Gemini and Meta Llama 3.1 exhibited lower average performance. Persistent challenges include the rigorous validation of bibliographic sources provided by AI models, which necessitates ongoing evaluation and cautious deployment, particularly in medical education and clinical practice.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the SBCOC for providing the exams and candidate results, which were essential for the completion of this study.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and substantially to the development of this article. CHK and KSPF contributed to the conceptualization of the study, identification and synthesis of the foundational literature, submission of questions to artificial intelligence models, data organization for subsequent analysis, and drafting and critical revision of the manuscript. LPS, MSTF, and RAZM contributed to the submission of questions to artificial intelligence models, data organization for subsequent analysis, and execution of the statistical analyses. MJST contributed to supervision, drafting, and critical revision of the manuscript.

DATA AVAILABILITY DECLARATION

The content underlying the research text is contained in the manuscript.

REFERENCES

1. Shortliffe EH, Buchanan BG. A model of inexact reasoning in medicine. Math Biosci. 1975;23(3-4):351-79. doi: 10.1016/0025-5564(75)90047-4.
2. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. N Engl J Med. 2018;378(11):981-3. doi: 10.1056/NEJMp1714229.
3. Universidade de São Paulo. Uso da IA no Instituto de Radiologia da USP permitirá laudos mais ágeis. Jornal da USP [Internet]. 2024 Feb 15 [cited 2025 Sep 16]. Available from: https://jornal.usp.br/radio-usp/uso-da-ia-no-instituto-de-radiologia-da-usp-permitira-laudos-mais-ageis/.
4. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89-94. doi: 10.1038/s41586-019-1799-6.
5. Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. Eur Heart J. 2019;40(25):2058-73. doi: 10.1093/eurheartj/ehz056.
6. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6(2):94-8. doi: 10.7861/futurehosp.6-2-94.
7. Paranjape K, Dybbro E, Kayembe C, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ. 2019;5(2):e16048. doi: 10.2196/16048.
8. Kung JE, Marshall C, Gauthier C, Gonzalez TA, Jackson JB 3rd. Evaluating ChatGPT performance on the orthopaedic in-training examination. JB JS Open Access. 2023;8(3):e23.00056. doi: 10.2106/JBJS.OA.23.00056.
9. Guerra GA, Hofmann H, Sobhani S, Hofmann G, Gomez D, Soroudi D, et al. GPT-4 artificial intelligence model outperforms ChatGPT, medical students, and neurosurgery residents on neurosurgery written board-like questions. World Neurosurg. 2023;179:e160-5. doi: 10.1016/j.wneu.2023.08.042.
10. Gilson A, Safranek C, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ. 2023;9(1):e45312. doi: 10.2196/45312.
11. Saab K, Tu T, Weng WH, Tanno R, Stutz D, Wulczyn E, et al. Capabilities of Gemini models in medicine. arXiv [preprint]. 2024 May 1. arXiv:2404.18416.
12. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230-43. doi: 10.1136/svn-2017-000101.
13. Carvalho ACPLF. Inteligência artificial: riscos, benefícios e uso responsável. Estudos Avançados. 2021;35(101):21-35. doi: 10.1590/s0103-4014.2021.35101.003.
14. Costa LPDB, Castro DHP, Cordeiro RP, Albino RB. Evaluation of the performance of ChatGPT/artificial intelligence in the multiple-choice test to obtain the title of specialist in orthopedics and traumatology. Acta Ortop Bras. 2025;33(spe1):e280947. doi: 10.1590/1413-785220243201e280947.

DOI: http://dx.doi.org/10.1590/1413-785220263404e300681

COMPARATIVE ANALYSIS OF OUTCOMES OF OPEN OR ENDOSCOPIC SPINAL CORD DECOMPRESSION: INTEGRATIVE LITERATURE REVIEW

ANÁLISE COMPARATIVA DOS DESFECHOS DA DESCOMPRESSÃO DA MEDULA ESPINHAL POR VIA ABERTA OU ENDOSCÓPICA: REVISÃO INTEGRATIVA DA LITERATURA

Victor Lima Dantas[1], *Rafael Quixabeira Bezerra de Araújo*[1], *Taynan Siqueira Massaro*[1], *Abraão Lucas Dias Clerot Muniz Paiva*[1]

1. Hospital Santa Marcelina de Itaquera, Serviço de Ortopedia e Traumatologia, Sao Paulo, SP, Brazil.

ABSTRACT

Introduction: Spinal cord compression is a neurological condition caused by pressure on the spinal cord, with etiologies including degenerative, neoplastic, infectious, traumatic, and malformative causes. It primarily affects the elderly, with symptoms ranging from pain and sensory-motor deficits to paralysis. The diagnosis is based on clinical evaluation and magnetic resonance imaging. Surgical treatment includes traditional open decompression and the endoscopic technique, the latter involving less trauma and quicker recovery. Objective: To review and compare the outcomes of spinal cord decompression via open and endoscopic approaches. Method: A search in the PUBMED database with the strategy: (decompre*[title] OR lamin*[title] OR discec*[title] OR micro*[title]) AND ((endosc*[title] OR minima*[title] OR percut*[title]) AND (open*[title] OR traditional[title])), with a 10-year time frame. Results: 22 studies were identified; after reviewing titles and abstracts, 11 were excluded, leaving 11 articles that formed the sample. Conclusion: Minimally invasive techniques showed clinical efficacy comparable to open approaches in the treatment of lumbar disc herniation and lumbar spinal stenosis, with advantages in perioperative parameters. Variations in outcomes related to pain, function, complications, and reoperations suggest that the choice of technique should be individualized, considering the clinical presentation, the complexity of the lesion, and surgical expertise. ***Level of evidence II; review article.***

Keywords: Minimally Invasive Surgical Procedures; Low Back Pain; Recovery of Function; Intraoperative Complications.

RESUMO

Introdução: A compressão medular é uma condição neurológica causada por pressão sobre a medula espinhal, com etiologias degenerativas, neoplásicas, infecciosas, traumáticas e malformativas. Afeta principalmente idosos, com sintomas variando de dor e déficit sensitivo-motor a paralisia. O diagnóstico baseia-se na avaliação clínica e na ressonância magnética. O tratamento cirúrgico inclui a descompressão aberta tradicional e a técnica endoscópica, esta com menor trauma e recuperação mais rápida. Objetivo: Revisar e comparar os desfechos da descompressão medular por via aberta e endoscópica. Método: Pesquisa na base PUBMED com a estratégia: (decompre[title] OR lamin*[title] OR discec*[title] OR micro*[title]) AND ((endosc*[title] OR minima*[title] OR percut*[title]) AND (open*[title] OR traditional[title])), com corte temporal de 10 anos. Resultados: Foram identificados 22 estudos; após leitura de títulos e resumos, 11 foram excluídos, restando 11 artigos que compuseram a amostra. Conclusão: As técnicas minimamente invasivas apresentaram eficácia clínica comparável às abordagens abertas no tratamento da hérnia de disco lombar e da estenose espinhal lombar, com vantagens em parâmetros perioperatórios. Variações nos desfechos de dor, função, complicações e reoperações sugerem que a escolha técnica deve ser individualizada, considerando o quadro clínico, a complexidade da lesão e a expertise cirúrgica.* ***Nível de evidência II; artigo de revisão.***

Descritores: *Procedimentos Cirúrgicos Minimamente Invasivos; Dor Lombar; Recuperação de Função Fisiológica; Complicações Intraoperatórias.*

Citation: Dantas VL, Araújo RQB, Massaro TS, Paiva ALDCM. Comparative analysis of outcomes of open or endoscopic spinal cord decompression: integrative literature review. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 7. Available from URL: http://www.scielo.br/aob.

All authors declare no potential conflict of interest related to this article.

The study was conducted at Hospital Santa Marcelina de Itaquera, Servico de Ortopedia e Traumatologia, Sao Paulo, SP, Brazil.
Correspondence: Victor Lima Dantas. 177, Rua Santa Marcelina, Itaquera, Sao Paulo, SP, Brazil. 08270-070. victor.limada13@gmail.com

Article received on 09/08/2025 approved on 01/15/2026.
Handling Editor: Mauro Emilio Conforto Gracitelli

<< SUMÁRIO

INTRODUCTION

Spinal cord compression is a clinical condition characterized by pressure exerted on the spinal cord by adjacent structures, resulting in neurological impairment of varying degrees [1]. This compression may have various etiologies, including degenerative diseases of the spine, such as spinal canal stenosis and large disc herniations; primary or metastatic neoplasms; infections such as epidural abscesses; trauma; and congenital malformations[2]. Furthermore, the progression of the condition may be acute, subacute, or chronic, depending on the underlying cause and the speed at which the compression occurs[1,2].

Spinal cord compression primarily affects elderly individuals due to the higher prevalence of degenerative pathologies of the spine in this age group. However, young adults may also be affected, especially in traumatic or tumoral contexts[2–4]. Studies indicate that symptomatic lumbar stenosis, for example, affects up to 30% of the population over 60 years of age[2,5,6]. Regarding the distribution by sex, some types of compression, such as traumatic ones, have a higher prevalence in men, while neoplastic and degenerative causes are more equitable[4,5].

The clinical consequences of spinal cord compression vary according to the level and severity of the injury, which may include intense pain, loss of muscle strength, impairment of tactile and proprioceptive sensitivity, alterations in sphincter control, and autonomic dysfunctions[1,7]. In severe cases, there may be partial or complete paralysis of the limbs below the level of the injury, directly impacting the patient's mobility and functional independence, as well as impairing their quality of life[2,7].

The diagnosis of spinal cord compression is based on the correlation between clinical data and imaging findings[8]. Additionally, a detailed neurological evaluation is essential, including a physical examination and investigation of motor, sensory, and autonomic signs[1]. Magnetic resonance imaging (MRI) of the spine is the primary imaging study used, allowing for precise identification of the level, extent, and cause of the compression. In specific situations, computed tomography (CT) and myelography may also be employed as complementary examinations[8].

The treatment of spinal cord compression may be conservative or surgical, depending on the etiology, severity of symptoms, and clinical condition of the patient[1]. Conservative measures include the use of analgesics, anti-inflammatories, physical therapy, and motor rehabilitation. However, in cases where there is progressive neurological deficit, intractable pain, or failure of clinical treatment, surgical intervention for spinal decompression becomes necessary, which may be performed through different access routes and surgical techniques[7,9].

Open spinal decompression is a traditional technique that involves direct access to the spine through a wide incision, with dissection of the paravertebral tissues and exposure of the spinal canal[10]. Procedures such as laminectomy, foraminotomy, or discectomy are performed to remove the structures causing compression, such as bone fragments or disc herniations[2,11]. Although effective, this approach is associated with greater tissue aggressiveness, longer recovery time, and risk of perioperative complications[9,11].

Endoscopic spinal decompression is a minimally invasive technique that allows for the removal of compressive elements through small incisions and the use of optical systems and specialized instruments[9]. This approach reduces trauma to soft tissues, preserves adjacent anatomical structures, and results in less blood loss, shorter hospitalization, and faster functional recovery, having gained traction as an effective alternative to the conventional technique in selected cases[2,11].

In light of the increasing adoption of minimally invasive approaches in spinal surgery, it becomes relevant to comparatively evaluate the outcomes of open and endoscopic spinal decompression techniques. Thus, this work focuses on conducting an integrative literature review to identify which of the surgical modalities presents better results in terms of efficacy, safety, and functional recovery, contributing to evidence-based clinical decision-making.

OBJECTIVE

To review the literature regarding the outcomes of spinal decompression via open or endoscopic methods and to compare them.

METHOD

This exploratory work was developed through an integrative literature review, aiming to synthesize the available evidence. The search was conducted in the PUBMED database, utilizing the following search strategy: *(decompre*[title] OR lamin*[title] OR discec*[title] OR micro*[title]) AND ((endosc*[title] OR minima*[title] OR percut*[title]) AND (open*[title] OR traditional[title]))*, considering a temporal cut-off of 10 years, from 2015 to 2025. Only clinical trials and observational studies were exclusively included.

The review was conducted according to the methodological steps described by Souza et al.[12], which comprise: (a) definition of the guiding question; (b) survey of the available scientific production; (c) initial screening of the selected studies; (d) critical evaluation of the articles by experts; (e) analysis and interpretation of the findings; and (f) integration of the similarities and divergences identified among the examined studies. The guiding question used was: "what are the comparative clinical outcomes between spinal decompression performed via open and endoscopic methods, according to the evidence available in the scientific literature of the last 10 years?"

RESULTS

Initially, 22 studies were identified that met the previously established search strategy. After reading the titles and abstracts, 11 articles were excluded for not directly addressing the outcomes of spinal decompression via open or endoscopic methods. The remaining 11 articles were read in full, summarized, and presented in the next section in chronological order considering the year of publication. Table 1 presents a summary of the main information related to the reviewed works.

DISCUSSION

Choi et al.[13], compared the results of lumbar disc herniation (LDH) treated with percutaneous endoscopic lumbar discectomy (PELD) and open lumbar microdiscectomy (OLM) through a retrospective observational study. The study was conducted from January 2011 to June 2012 with 44 consecutive patients diagnosed with LDH without cauda equina syndrome, scheduled for spinal surgery. Clinical outcomes were assessed using a visual analog scale (VAS), functional status was evaluated by the Oswestry Disability Index (ODI) at one, six, and 24 months postoperatively, in addition to the surgical satisfaction rate at the final follow-up. The radiological variables were evaluated through simple radiography. Forty-three patients were included; 20 and 23 patients underwent PELD and OLM, respectively. Both groups exhibited significant improvements in leg and back pain postoperatively, and although there was no significant difference in this regard between the groups, the improvement in back pain was significantly greater in the PELD group than in the OLM group. The surgical satisfaction rate of the PELD group was significantly higher than that of the OLM group, and the average operation time, hospital stay, and time until return to work were significantly shorter in the PELD group than in the OLM group. The height of the disc (%) significantly decreased from 23.7 ± 3.3

Table 1. Summary of the main information related to the reviewed works.

Authors	Title	Type of Study	Number of patients	Conclusion
Choi et al. [13]	*Percutaneous Endoscopic Lumbar Discectomy as an Alternative to Open Lumbar Microdiscectomy for Large Lumbar Disc Herniation*	Retrospective observational study	43	Percutaneous endoscopic lumbar discectomy (PELD) was effective in the treatment of lumbar disc herniation (LDH), with advantages such as reduced lumbar pain, preservation of disc height, faster recovery, and greater satisfaction compared to open lumbar microdiscectomy (OLM).
Cristante et al. [14]	*Randomized clinical trial comparing lumbar percutaneous hydrodiscectomy with lumbar open microdiscectomy for the treatment of lumbar disc protrusions and herniations*	Randomized clinical trial	40	Hydrodiscectomy was as effective as open microdiscectomy, with similar rates of complications, recurrence, and satisfaction.
Chen et al. [15]	*Traditional versus Percutaneous Transforaminal Endoforaminal Discectomy Effect on Nervous System Function and Serum LEK, GFAP, and PGE-2 in Patients with Senile Lumbar Spinal Stenosis*	Comparative experimental study	146	Percutaneous transforaminal endoscopic discectomy (PTED) proved superior to traditional surgery in terms of less blood loss, shorter hospitalization, and better neurological and lumbar function.
Ahn et al. [16]	*Transforaminal Endoscopic Lumbar Discectomy Versus Open Lumbar Microdiscectomy: A Comparative Cohort Study with a 5-Year Follow-Up*	Prospective study with retrospective analysis	298	Transforaminal endoscopic lumbar discectomy (TELD) had clinical outcomes comparable to OLM, with advantages such as shorter surgical time, hospitalization, and return to work, despite the lack of randomization.
Song et al. [17]	*Comparison of the Outcomes of Percutaneous Endoscopic Interlaminar Lumbar Discectomy and Open Lumbar Microdiscectomy at the L5-S1 Level*	Observational, retrospective, and matched cohort study	56	Percutaneous interlaminar endoscopic discectomy (PIED) had better perioperative results (less pain, shorter hospitalization and surgery time), although medium-term clinical and radiological outcomes were similar to those of OLM.
Gadjradj et al. [18]	*Full endoscopic versus open discectomy for sciatica: randomized controlled non-inferiority trial*	Multicenter randomized clinical trial	488	PTED was not inferior to OLM in reducing leg pain and had slightly more favorable results in pain, function, and recovery, thus can be considered an effective alternative.
Quian et al. [19]	*Transforaminal endoscopic lumbar discectomy using a 45□ puncture angle and foraminotomy versus traditional THESYS for L5/S1 lumbar disc herniation: a prospective randomized controlled trial*	Randomized prospective clinical trial	Not Reported	The TELD at 45° was superior to traditional Thomas Hoogland spinal endoscopy systems (THESYS) in terms of operative time, pain (visual analog scale - VAS), Oswestry Disability Index (ODI), and lower complication rates; after three months, the outcomes were similar.
Baranidharan et al. [20]	*24-Month Outcomes of Indirect Decompression Using a Minimally Invasive Interspinous Fixation Device versus Standard Open Direct Decompression for Lumbar Spinal Stenosis: A Prospective Comparison*	Randomized, prospective, multicenter clinical trial	48	The interspinous fixation device (IFD) proved to be safe and effective over 2 years, with significant improvement in pain, function, and less blood loss compared to surgical decompression.
Seddighi et al. [21]	*Clinical Outcomes at 2-Year Follow-Up Comparing Open Surgery and Percutaneous Laser Disc Decompression for Radicular Sciatic Pain Patients*	Randomized and controlled prospective clinical trial	84	Percutaneous laser disc decompression (PLDD) and open surgery had similar efficacy in pain and function in the long term; PLDD had a higher reoperation rate but remains a viable alternative.
Hao et al. [22]	*Clinical outcomes of unilateral biportal endoscopic discectomy (UBE) compared with conventional open lumbar discectomy with 3D microscope (OLDM) assisted*	Comparative and retrospective observational study	76	The unilateral biportal endoscopic technique (UBE) was comparable to open lumbar discectomy with a 3D microscope (OLDM) in clinical outcomes, but had advantages in shorter operative time, less blood loss, and shorter hospitalization.
Broekema et al. [18]	*Percutaneous pedicle screw placement with a mini-open decompression versus open surgery in the treatment of lumbar spondylolisthesis: one-year results of a randomised controlled trial*	Multicenter randomized clinical trial	169	There was no significant difference between mini-open surgery and open surgery for spondylolisthesis; the expected advantages of mini-open surgery were not confirmed.

Source: Data collected by the authors.

to 19.1 ± 3.7 after OLM, but did not change significantly after PELD (23.6 ± 3.2 to 23.4 ± 4.2). The segmental angle at the operated level increased from 10.3° to 15.4° in the PELD group, which was significantly greater than in the OLM group (9.6° to 11.6°). In the OLM group, there was one case of fusion due to instability, and in the PELD group, one case required revision surgery and another presented recurrence. There were no perioperative complications in either group. According to the authors, despite the small sample size and short follow-up period, it was possible to observe that PELD was an effective treatment for LDH and was associated with potential advantages, including rapid recovery, improvement of back pain, and preservation of disc height.

Cristante and collaborators[14] evaluated the results of hydrodiscectomy compared to open microdiscectomy concerning pain, function, satisfaction, complications, and recurrence rates. To this end, they conducted a randomized clinical trial with patients referred to the authors' service for low back pain. Such subjects were included in the study if they presented disc protrusion or a small hernia at only one level, without neurological deficits and without resolution after six weeks of conservative treatment. One group underwent open microdiscectomy and the other group underwent percutaneous microdiscectomy by hydro-surgery. Function was assessed by the ODI and pain was assessed by a VAS. The evaluations were conducted preoperatively and subsequently during the first week and after one, three, six, and twelve months postoperatively. During the study period, 20 patients were included in each arm, and 39 completed one year of follow-up (one patient died from unrelated causes). Both groups showed equal improvement in the VAS and ODI assessments after treatment, with no significant differences. Furthermore, the improvement in the lumbar VAS score was not significant in the hydrodiscectomy group. The rates of infection, pain, recurrence, and satisfaction were similar between the two groups. According to the authors, percutaneous hydrodiscectomy proved to be as effective as open microdiscectomy in reducing pain. The rates of complications and recurrence of hernia were similar between the groups, and the patients' satisfaction with the treatment was also comparable.

Chen et al.[15], compared the influence of percutaneous transforaminal endoscopic discectomy (PTED) and traditional surgery on the function of the nervous system and the serum levels of leuencephaline (LEK), glial fibrillary acidic protein (GFAP), and prostaglandin E-2 (PGE-2) in patients with senile lumbar spinal stenosis. To this end, they conducted an experimental study from March 2017 to March 2018 with 146 patients suffering from senile lumbar spinal stenosis from a single clinical center. The subjects were randomly divided into a control group and an observation group, with 73 in each group. The control group underwent traditional surgery, while the observation group underwent PTED. They compared the overall situation of the operation, serum LEK, GFAP, PGE-2, the score of the *American Spinal Injury Association* (ASIA), and the score of the *Japanese Orthopaedic Association* (JOA). The intraoperative blood loss in the observation group was less than in the control group; moreover, both the operative time and the duration of hospitalization in the observation group were shorter than in the control group. After 24 hours post-operation, the serum levels of LEK and the ASIA score in the observation group were higher than in the control group, while the serum levels of GFAP and PGE-2 and the JOA score in the observation group were lower than in the control group. According to the authors, compared to traditional surgery, PTED presented the advantages of less intraoperative blood loss, shorter operative time, and reduced duration of hospitalization, among others. Furthermore, PTED would effectively reduce the serum expression of LEK, BFGF, and PGE-2 in patients, and drastically improve the function of the nervous system and lumbar function.

Ahn and collaborators[16], demonstrated the clinical outcomes of transforaminal endoscopic lumbar discectomy (TELD) compared to those of open lumbar microdiscectomy (OLM). To this end, between January 2009 and September 2011, 335 consecutive patients with symptomatic herniated discs were treated with decompressive discectomy, either by TELD or by open microdiscectomy. The patients were prospectively entered into the clinical database, and their medical records were reviewed retrospectively. Data from 298 patients treated with decompressive discectomy, either by TELD or by OLM, were evaluated with a minimum follow-up period of five years. Among them, 146 patients were treated with TELD (TELD group) and the remaining 152 patients with open microdiscectomy (OLM group). Perioperative data and clinical outcomes were assessed using the Visual Analog Scale (VAS), Oswestry Disability Index (ODI), and modified Macnab criteria. The VAS and ODI improved significantly in both groups, with an excellent or good outcome rate of 88.36% and 87.5% in the TELD and OLM groups, respectively. The reoperation rate was 4.2% and 3.3% in the TELD and OLM groups, respectively, and there were no significant differences in clinical outcomes. However, the operative time, length of hospital stay, and time to return to work were significantly shorter in the TELD group. Regarding the limitations of the study, the authors noted that the selection of patients was not randomized; therefore, the risk of bias may be increased. Furthermore, the study lacked an analysis of the radiographic changes related to degenerative alteration over the long-term follow-up period. The researchers concluded that the long-term results of TELD for soft LDH were comparable to those of conventional open microdiscectomy. Moreover, the technique of selective endoscopic discectomy under local anesthesia offered the typical advantages of minimally invasive procedures, such as reduced operative time, hospital stay, and recovery time.

Song et al.[17], compared the clinical, surgical, and radiological outcomes of patients with disc herniation at the L5-S1 level who underwent endoscopic percutaneous interlaminar discectomy (PEID) or OLM, performed by a single surgeon with beginner-level proficiency. To this end, they conducted a retrospective observational study with a matched cohort that included 56 patients who underwent discectomy at the L5-S1 level, with a minimum follow-up of one year, between September 2012 and August 2016. The patients were allocated into two groups: a PEID group (n = 27; September 2014 to August 2016) and an OLM group (n = 29; September 2012 to August 2014). The clinical, surgical, and radiological outcomes were evaluated retrospectively. The baseline characteristics, including age, sex, medical history, body mass index, preoperative symptoms, and preoperative radiological findings, did not differ significantly between the groups. Furthermore, the overall clinical outcomes, including back and leg pain; the surgical outcomes, including blood loss, complication rate, and recurrence rate; and the radiological results, including degree of decompression, disc height, and sagittal alignment, were not significantly different between the two groups. However, the PEID group demonstrated significant advantages, including less immediate postoperative low back pain, favorable immediate postoperative Odom criteria, shorter operation time, shorter hospital stay, and a quicker return to work. According to the authors, their findings indicated that the PEID group achieved better perioperative outcomes, despite no significant difference between the groups in clinical and radiological results at mid-term follow-up.

Gadjradj and collaborators[18], assessed whether DET was non-inferior to OLM in reducing leg pain caused by LDH. To this end, they conducted a multicenter randomized clinical trial with a non-inferiority design in four hospitals involving 613 patients aged between 18 and 70 years with at least six weeks of radiating leg pain caused by lumbar disc herniation. The study included

a predetermined set of 125 patients who underwent DET, which were the cases with a learning curve performed by surgeons who had not performed DET prior to the study. Thus, the patients were separated into two groups, DET (n = 179) and OLM (n = 309). The primary outcome was self-reported leg pain, measured by a VAS from 0 to 100 at 12 months, assuming a non-inferiority margin of 5.0. The secondary outcomes included complications, reoperations, self-reported functional status, measured by the ODI, VAS for back pain, health-related quality of life, and self-perceived recovery. The outcomes were measured up to one year after surgery and analyzed longitudinally according to the intention-to-treat principle, and patients belonging to the DET learning curve were omitted from the primary analyses. At 12 months, patients randomized to DET reported a significantly lower VAS score for leg pain compared to patients randomized to open microdiscectomy. Blood loss was lower, the duration of hospital stay was shorter, and the timing of postoperative mobilization was earlier in the DET group than in the OLM group. The secondary outcomes reported by the patients, such as the ODI, the VAS for back pain, health-related quality of life, and self-perceived recovery, were similarly favorable to the DET. In one year, nine (5%) in the DET group, compared to 14 (6%) in the OLM group, underwent repeat surgery. The per-protocol analysis and sensitivity analyses, including patients from the learning curve, resulted in outcomes similar to the primary analysis. According to the authors, the DET was not inferior to open microdiscectomy in reducing leg pain. The DET showed more favorable results for self-reported leg pain, back pain, functional status, quality of life, and recovery. These differences, however, were small and may not reach clinical relevance; nevertheless, the DET could be considered an effective alternative to OLM in the treatment of sciatica pain.

Qian et al.[19], conducted a prospective comparison of the efficacy and safety of TELD with a puncture angle of 45° *versus* the traditional Thomas Hoogland spinal endoscopy systems (THESYS) for the surgical treatment of LDH at L5/S1. To this end, consecutive patients with LDH L5/S1 undergoing TELD were randomized (1:1) and allocated to the TELD at 45° or THESYS groups. Clinical outcomes were assessed preoperatively, on the first day, and at three and six months post-surgery until the final follow-up. Surgical-related parameters, VAS scores, ODI, and modified MacNab criteria, as well as surgical complications, were recorded and analyzed. All patients were followed for at least 24 months, and compared to the THESYS group, the TELD at 45° group exhibited shorter operative time and intraoperative radiation time, as well as lower VAS scores for lumbar pain and leg pain during the intraoperative period. The VAS and ODI in the TELD at 45° group were significantly better than in the THESYS group at three months postoperatively. However, from three months onward, both groups exhibited comparable VAS and ODI scores. There was no significant difference between the two groups according to the modified MacNab criteria. There were two cases of residual disc and two cases of recurrence that required reoperation in the THESYS group. According to the authors, in cases of LDH L5/S1, the TELD technique at 45° was superior to the traditional THESYS in terms of surgical-related parameters and faster improvement in VAS and ODI, with a lower complication rate.

Baranidharan and collaborators[20], conducted an early-stage, multicenter, randomized clinical trial with a five-year follow-up comparing the efficacy of a minimally invasive interspinous fixation device (IFD), implanted laterally, for direct open surgical decompression in the treatment of lumbar spinal stenosis (LSS). The results of the two-year study were presented for 48 participants randomly assigned to IFD or decompression. The primary outcomes of the study included changes from baseline at the eight-week, six, 12, and 24-month follow-ups for leg pain via VAS, back pain (VAS), ODI, physical function of LSS (via Zurich Claudication Questionnaire), distance walked in five minutes, and the number of repetitions from sitting to standing in one minute. The secondary outcomes of the study included the global impression of the patient and physician regarding change, adverse events, reoperations, operative parameters, and fusion rate. Both treatment groups demonstrated statistically significant improvements in average leg pain, back pain, ODI, physical function of the ESL, distance walked, and repetitions seated and standing compared to baseline over 24 months. The average reduction in ODI compared to baseline levels was between 35% and 56% for IFD and 49% to 55% for decompression at all follow-up time points. The average reduction in leg pain for the IFD group was between 57% and 78% for all time points, with 72% to 94% of participants experiencing at least a 30% reduction in leg pain from eight weeks to 24 months. The distance walked for the IFD group increased from 66% to 94%, and the repetitions seated and standing increased from 44% to 64% for all follow-up time points. Blood loss was 88% lower in the IFD group, and the operative time parameters strongly favored IFD compared to decompression. A fusion rate of 89% was assessed in a subset of IFD participants, and there were no issues with the intraoperative device or reoperations in the IFD group; only one consolidated and asymptomatic spinous process fracture was observed at 24 months. For the authors, despite the low number of participants in the IFD group, the study demonstrated successful clinical and safety outcomes over two years for IFD, with significant advantages related to the operation compared to surgical decompression.

Seddighi et al.[21], compared the clinical outcomes of percutaneous laser disc decompression (PLDD) and open surgery in patients with radicular sciatic pain caused by LDH over a two-year follow-up period. To this end, they conducted a randomized controlled prospective clinical trial with 84 patients with chronic radicular pain allocated to the open surgery group (n = 42) or the PLDD group (n = 42). Patients were evaluated at the beginning of the study and at four, eight, 24, 48, and 96 weeks after the intervention. Outcome measures included the Roland-Morris Disability Questionnaire, the VAS for leg and back pain, and the body pain and physical functioning subscales of the *Short Form*-36. Reoperation rates were also recorded. No significant differences were observed in the Roland-Morris Disability Questionnaire scores, the VAS for leg and back pain, or the *Short Form*-36 between the two groups at any follow-up time point. Both groups showed improvement in disability and pain scores over time, with similar recovery patterns. The median reoperation rates were 19.0% for open surgery and 31.0% for PLDD (P = 0.314), indicating comparable long-term efficacy of both treatments. For the authors, the study demonstrated that PLDD and open surgery provide similar long-term outcomes in terms of disability, pain relief, and physical function for patients with radicular sciatic pain. They also noted that although PLDD was associated with a higher reoperation rate, it remained a viable minimally invasive alternative to open surgery. However, further research was necessary to refine patient selection criteria and improve the efficacy of the procedure for both interventions.

Hao and collaborators[22], evaluated and compared perioperative parameters and clinical outcomes, including operative time, intraoperative blood loss, pain and modification, hospital stay, patient satisfaction, and complications, between open lumbar discectomy with a 3D microscope (OLDM) and the unilateral biportal endoscopic technique (UBE) for LDH. A total of 76 patients with LDH were included in this study from February 2019 to February 2022. All of them had undergone spinal surgery and were subjected to OLDM (42 cases) and UBE (34 cases) in two distinct surgical centers. Respectively, all patients had level 1 lumbar disc herniation. The estimated blood loss, operation time, hospital stay duration, and patient complications were compared between the two groups.

The VAS for back and leg pain, the ODI, and the modified MacNab criteria were tested before surgery and three days, three months, and 12 months after surgery. Compared to the OLDM group, the UBE group had a shorter operative time, less intraoperative blood loss, and a shorter hospital stay. Furthermore, the VAS and ODI scores showed a significant reduction in both groups after the operation. There was no significant difference in the VAS and ODI scores preoperatively and three days, three months, and 12 months after the operation between the two groups. Meanwhile, there was no significant difference in the operational conversion rate and complications between the two groups. According to the authors, the application of OLDM produced clinical results similar to UBE for the treatment of LDH, including pain control and patient satisfaction. However, UBE was associated with several advantages over OLDM in terms of surgical time, intraoperative blood loss, short-term postoperative pain relief, and postoperative hospitalization.
Finally, Broekema et al.[23] compared the placement of percutaneous pedicle screws *versus* open screws in patients undergoing decompression of the mid-lumbar line due to symptomatic lumbar spondylolisthesis, focusing on short-term low back pain. To this end, they conducted a randomized clinical trial carried out in surgical centers between 2015 and 2020. Participants with spondylolytic or degenerative lumbar spondylolisthesis were randomized to receive percutaneous pedicle screw placement with mini-open decompression (mini-open surgery) or conventional open surgery with instrumented fusion (open surgery). The primary outcome was short-term low back pain after two weeks, measured by an VAS. Leg pain, disability, and quality of life were recorded at two and six weeks, three and six months, and one year. Surgical variables, including complications, were recorded, and analyses were performed on the intention-to-treat population. In total, 169 participants were included and randomized to mini-open surgery (n = 81) or standard open surgery (n = 88). No statistically or clinically significant differences were found between the groups in terms of primary or secondary outcomes, and the duration of surgery, blood loss, hospital stay, and complications were also similar between the groups. According to the authors, the study did not detect a difference in outcomes between mini-open surgery compared to open surgery in patients with spondylolisthesis. Thus, the hypothetical advantage of short-term reduction of lower back pain, less blood loss, and better clinical outcomes could not be confirmed.

The analyzed studies largely agree on the clinical efficacy of various minimally invasive techniques for the treatment of lumbar disc herniation and spinal stenosis when compared to traditional open surgical approaches. Most studies reported significant improvement in VAS, ODI, and patient satisfaction in both groups, with few relevant clinical differences in the long term. There is consensus that endoscopic or percutaneous techniques, such as PELD, TELD, DET, PEID, hydrodiscectomy, and UBE, provided perioperative advantages, including shorter surgical time, less blood loss, shorter hospitalization, and quicker return to activities, as highlighted in studies such as those by Choi et al.[13], Ahn et al.[16], Song et al.[17], Hao et al.[22], and Gadjradj et al.[18]. However, some discrepancies emerge: while some studies, such as that by Cristante et al.[14], did not observe significant superiority of the minimally invasive technique over open microdiscectomy in terms of pain and function, others, such as that by Chen et al.[15], reported additional neurophysiological benefits with endoscopic techniques. There was also variation in complication and reoperation rates, with studies such as that by Seddighi et al.[21], suggesting a higher reoperation rate with percutaneous techniques, although without significant clinical impact. Finally, Broekema et al.[23], e Baranidharan et al.[20], drew attention to the absence of clear clinical superiority between open and minimally invasive techniques in specific cases, emphasizing the importance of appropriate patient selection. Thus, despite the predominance of operative advantages of minimally invasive techniques, the findings are not uniform across all clinical contexts, and the choice of method should consider factors such as the type and level of the lesion, the surgeon's experience, and the patient's profile.

CONCLUSION

The analyzed studies indicated that minimally invasive techniques, such as endoscopic and percutaneous discectomies, demonstrated clinical efficacy comparable to open surgical approaches in the treatment of lumbar disc herniation and lumbar spinal stenosis, with consistent advantages in perioperative parameters. However, the evidence also demonstrated variations in outcomes related to pain, function, complication rates, and reoperations, suggesting that the choice of the most appropriate technique should be individualized, considering the clinical picture, the complexity of the lesion, and the expertise of the surgical team.

CONTRIBUTIONS OF THE AUTHORS

Each author made a personal and significant contribution to the development of this article. DVL and ARQB primarily engaged in the selection and evaluation of the reviewed articles, in the writing, and secondarily in the final review of the submitted manuscript. MTS and PALDCM primarily contributed to the writing of the manuscript and the final review of the submitted article, and secondarily in the selection and evaluation of the selected articles.

DATA AVAILABILITY DECLARATION

the underlying contents of the research text are contained in the manuscript.

REFERENCES

1. Ko HY. A Practical Guide to Care of Spinal Cord Injuries: Clinical Questions and Answers. Springer Nature; 2023. 832 p.
2. Rajendram R, Preedy VR, Martin CR. Diagnosis and Treatment of Spinal Cord Injury. Academic Press; 2022. 628 p.
3. Jaroudi S, Faiz SA, Fuller G, Rubio D. Spinal cord compression revealing metastatic cancer. Am J Med Sci. 2022;363(6):e56-7. doi: 10.1016/j.amjms.2021.10.011.
4. Van den Brande R, MJ Cornips E, Peeters M, Ost P, Billiet C, Van de Kelft E. Epidemiology of spinal metastases, metastatic epidural spinal cord compression and pathologic vertebral compression fractures in patients with solid tumors: A systematic review. J Bone Oncol. 2022;35:100446. doi: 10.1016/j.jbo.2022.100446.
5. Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and Management of Lumbar Spinal Stenosis: A Review. JAMA. 2022;327(17):1688-99. doi: 10.1001/jama.2022.5921.
6. Farahbakhsh F, Khosravi S, Baigi V, Pourghahramani Koltapeh M, Khayyamfar A, Eskandari Z, et al. The Prevalence of Asymptomatic Cervical Spinal Cord Compression in Individuals Presenting With Symptomatic Lumbar Spinal Stenosis: A Meta-Analysis. Glob Spine J. 2024;14(3):1052-60. doi: 10.1177/21925682231202776.
7. Esperança-Martins M, Roque D, Barroso T, Abrunhosa-Branquinho A, Belo D, Simas N, et al. Multidisciplinary Approach to Spinal Metastases and Metastatic Spinal Cord Compression—A New Integrative Flowchart for Patient Management. Cancers (Basel). 2023;15(6):1796. doi: 10.3390/cancers15061796.
8. Kuah T, Vellayappan BA, Makmur A, Nair S, Song J, Tan JH, et al. State-of-the-Art

Imaging Techniques in Metastatic Spinal Cord Compression. Cancers (Basel). 2022;14(13):3289. doi: 10.3390/cancers14133289.
9. Meyer HS, Wagner A, Raufer A, Joerger AK, Gempt J, Meyer B. Surgery in Acute Metastatic Spinal Cord Compression: Timing and Functional Outcome. Cancers (Basel). 2022;14(9):2249. doi: 10.3390/cancers14092249.
10. Echt M, Stock A, Ramos RD la G, Der E, Hamad M, Holland R, et al. Separation surgery for metastatic epidural spinal cord compression: comparison of a minimally invasive versus open approach. Neurosurg Focus. 2021;50(5):E10. doi: 10.3171/2021.2.FOCUS201124.
11. Alshareef M, Klapthor G, Alawieh A, Lowe S, Frankel B. Evaluation of open and minimally invasive spinal surgery for the treatment of thoracolumbar metastatic epidural spinal cord compression: a systematic review. Eur Spine J. 2021;30(10):2906-14. doi: 10.1007/s00586-021-06880-7.
12. Souza MT de, Silva MD da, Carvalho R de. Revisão integrativa: o que é e como fazer. Einstein São Paulo. 2010;8(1 Pt 1):102-6. doi: 10.1590/S1679-45082010RW1134.
13. Choi KC, Kim JS, Park CK. Percutaneous Endoscopic Lumbar Discectomy as an Alternative to Open Lumbar Microdiscectomy for Large Lumbar Disc Herniation. Pain Physician. 2016;19(2):E291-300.
14. Cristante AF, Rocha ID, Marcon RM, Filho TEPB. Randomized clinical trial comparing lumbar percutaneous hydrodiscectomy with lumbar open microdiscectomy for the treatment of lumbar disc protrusions and herniations. Clin Sao Paulo Braz. 2016;71(5):276-80. doi: 10.6061/clinics/2016(05)06.
15. Chen H, Xue J, Wu X. Traditional versus Percutaneous Transforaminal Endoforaminal Discectomy Effect on the Nervous System Function and Serum LEK, GFAP and PGE-2 in Patients with Senile Lumbar Spinal Stenosis. J Coll Physicians Surg Pak. 2019;29(4):341-4. doi: 10.29271/jcpsp.2019.04.341.
16. Ahn Y, Lee SG, Son S, Keum HJ. Transforaminal Endoscopic Lumbar Discectomy Versus Open Lumbar Microdiscectomy: A Comparative Cohort Study with a 5-Year Follow-Up. Pain Physician. 2019;22(3):295-304.
17. Song SK, Son S, Choi SW, Kim HK. Comparison of the Outcomes of Percutaneous Endoscopic Interlaminar Lumbar Discectomy and Open Lumbar Microdiscectomy at the L5-S1 Level. Pain Physician. 2021;24(4):E467-75.
18. Gadjradj PS, Rubinstein SM, Peul WC, Depauw PR, Vleggeert-Lankamp CL, Seiger A, et al. Full endoscopic versus open discectomy for sciatica: randomised controlled non-inferiority trial. BMJ. 2022;376:e065846. doi: 10.1136/bmj-2021-065846.
19. Qian Y, Chen Z, Zhang G, Wang X, Liu J, Yang A, et al. Transforaminal endoscopic lumbar discectomy using a 45° puncture angle and foraminotomy versus traditional THESYS for L5/S1 lumbar disc herniation: a prospective randomized controlled trial. Eur Spine J. 2024;33(2):453-62. doi: 10.1007/s00586-023-08117-1.
20. Baranidharan G, Bretherton B, Feltbower RG, Timothy J, Khan AL, Subramanian A, et al. 24-Month Outcomes of Indirect Decompression Using a Minimally Invasive Interspinous Fixation Device versus Standard Open Direct Decompression for Lumbar Spinal Stenosis: A Prospective Comparison. J Pain Res. 2024;17:2079-97. doi: 10.2147/JPR.S453343.
21. Seddighi AS, Seddighi A, Hosseini S. Clinical Outcomes at 2-Year Follow-Up Comparing Open Surgery and Percutaneous Laser Disc Decompression for Radicular Sciatic Pain Patients. World Neurosurg. 2025;194:123392. doi: 10.1016/j.wneu.2024.10.121.
22. Hao J, Chen R, Zheng J, Xu S, Xue H, Yao Y. Clinical outcomes of unilateral biportal endoscopic discectomy (UBE) compared with conventional open lumbar discectomy with 3D microscope (OLDM) assisted. Medicine (Baltimore). 2025;104(6):e41440. doi: 10.1097/MD.0000000000041440.
23. Broekema AEH, Schenck CD, Wolfs JFC, Torensma B, de Ruiter GCW, Kuijlen JMA, et al. Percutaneous pedicle screw placement with a mini-open decompression versus open surgery in the treatment of lumbar spondylolisthesis: one-year results of a randomised controlled trial. Eur Spine J. 2025;34(6):2407-15. doi: 10.1007/s00586-025-08855-4.

DOI: http://dx.doi.org/10.1590/1413-785220263404e295387

PRP INFILTRATION *VERSUS* CORTICOSTEROIDS: A PRELIMINARY RANDOMIZED CLINICAL TRIAL

INFILTRAÇÃO COM PRP *VERSUS* CORTICOSTEROIDES: ENSAIO CLÍNICO RANDOMIZADO PRELIMINAR

William Zarza Santos[1], *Roberta Sessa Stilhano Yamaguchi*[1], *Edilson Silva Machado*[1], *Rodrigo Góes Medéa de Mendonça*[1], *Alberto Gotfryd*[1], *Robert Meves*[1]

1. Irmandade da Santa Casa de Misericordia de Sao Paulo, Departamento de Ortopedia e Traumatologia, Grupo de Afeccoes da Coluna Vertebral, Sao Paulo, SP, Brazil.

ABSTRACT

Introduction: Chronic low back pain is one of the most prevalent health conditions, with significant social and economic impact. Its increasing incidence has driven the demand for new, effective therapeutic alternatives. This study compared the efficacy of platelet-rich plasma (PRP) and corticosteroid injections in the treatment of chronic low back pain of facet joint origin. Methods: This was a randomized clinical trial involving patients with chronic low back pain of facet joint origin, selected through a positive response to medial branch block. Epidemiological data and clinical outcomes were analyzed, including disability questionnaires (Roland-Morris Disability Questionnaire) and the visual analog scale (VAS) for pain. Assessments were conducted at baseline, one month, and three months post-procedure. Results: A total of 59 patients were included, with 32 in the corticosteroid group and 27 in the PRP group. No significant differences were found between the treatments in terms of disability scores or the Roland-Morris Questionnaire. However, a significant interaction between treatment and time was observed (p = 0.038). Regarding pain scores, no significant differences were found between groups at any of the assessed time points. Conclusion: The study found no evidence of superiority of platelet-rich plasma over corticosteroid injections in terms of pain reduction and functional outcomes for the treatment of lumbar facet syndrome. ***Level of evidence: I; Randomized clinical trial.***

Keywords: Platelet-Rich Plasma; Low Back Pain; Injections; Adrenal Cortex Hormones.

RESUMO

Introdução: A dor lombar crônica é uma das condições de saúde mais prevalentes, com significativo impacto social e econômico. A sua crescente incidência tem gerado a necessidade de novas opções terapêuticas eficazes. Este estudo comparou a eficácia das infiltrações de plasma rico em plaquetas (PRP) e corticosteroides no tratamento da dor lombar crônica de origem facetária. Método: Trata-se de um ensaio clínico randomizado que avaliou pacientes com dor lombar crônica de origem facetária, selecionados por resposta positiva ao bloqueio do ramo medial, em relação aos efeitos dos tratamentos. Foram analisados dados epidemiológicos e desfechos clínicos, incluindo o Questionário de Incapacidade Roland-Morris e a escala visual analógica (EVA) de dor, com avaliações antes, 1 mês e 3 meses após o procedimento. Resultados: Foram incluídos 59 pacientes, 32 no grupo corticoide e 27 no grupo PRP. Não houve diferença significativa entre os tratamentos nos questionários de incapacidade e Roland Morris, embora tenha sido observada uma interação significativa entre tratamento e tempo (p=0,038). Na escala de dor, não houve diferença entre os tratamentos em nenhum dos tempos avaliados. Conclusão: O estudo evidenciou ausência de superioridade do Plasma Rico em Plaquetas em comparação ao corticoide na avaliação de dor e escores de função no tratamento da síndrome facetária lombar. ***Nível de evidência I; Estudo clínico randomizado.***

Descritores: *Plasma Rico em Plaquetas; Dor Lombar; Injeções; Corticosteroides.*

Citation: Santos WZ, Yamaguchi RSS, Machado ES, Mendonça RGM, Gotfryd A, Meves R. PRP infiltration *versus* corticosteroids: A preliminary randomized clinical trial. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 5. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

Low back pain is one of the most common health problems, with high prevalence and social cost[1]. Although we attempt to define low back pain based on lumbar structures (discogenic pain, facet joint pain, sacroiliac joint pain, myofascial pain), diagnostic investigations have a limited role. Both clinical tests show insufficient accuracy in this differentiation, as do imaging methods, which, although they reveal changes, often do not have clinical correlation with the type of pain[2].

In light of the increasing cost, incidence, and prevalence of people with chronic low back pain, better treatment options have become a major point of discussion[3]. Treatment modalities such as corticosteroid injections have gained significant traction; however, there are limitations to this treatment regarding the frequency and

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Irmandade da Santa Casa de Misericordia de Sao Paulo.
Correspondence: William Zarza Santos. 112, Rua Dr. Cesario Motta Junior, Vila Buarque, Sao Paulo, SP, Brazil. 01221-020. w.zarzasantos@gmail.com.

Article received on 03/28/2025 approved on 04/02/2026.
Handling Editor: Eduardo Angeli Malavolta

duration of effect, as well as potential toxic properties to tendons and cartilage associated with these injections[4].

In this context of high prevalence of low back pain, combined with the scarcity of effective treatment for the problem, the need for a therapeutic alternative with good clinical results arises, which presents lower morbidity and is supported by clinical trials[3]. This is precisely where biological therapies emerge as a promising alternative, with Platelet-Rich Plasma (PRP) and mesenchymal stem cells (MSCs) being the most prominent orthobiologics in the treatment of musculoskeletal pain.

The knowledge that platelets carry various growth factors and tissue regeneration factors led to the hypothesis that the application of a platelet concentrate could result in modulation of inflammation as well as a regenerative stimulus in an injured or degenerated tissue[5].

In the treatment of low back pain, the literature suggests benefits in the use of PRP[6], however, what we still observe is a significant limitation of randomized clinical trials, especially related to facet joint pain, since most studies focus on intradiscal PRP injection aimed at discogenic low back pain[7].

In addition to the scarcity of literature on randomized controlled clinical trials among patients with facet joint pain, it is also a fact that PRP is not yet a therapeutic method recommended by Brazilian medical authorities. Such factors increase the importance and relevance of the current work.

The objective of this work is to compare the efficacy of PRP with the standard treatment for facet joint pain: corticosteroid injection.

MATERIALS AND METHODS

A double-blind randomized clinical trial was conducted, comparing the injection of platelet-rich plasma with the standard treatment (corticosteroid injection).

Patients over 18 years old with low back pain for at least 6 months and radiographs confirming degenerative changes in the lumbar spine were recruited for the study. Pain with a score of five or higher on the Visual Analog Scale (VAS) was also a criterion for recruitment in the study.

The following exclusion criteria were applied: pain radiating to the limbs (sciatica), previous lumbar spine surgery, radiographs showing spondylolisthesis, fractures, discitis, or tumors, pregnancy, history of drug abuse, allergy to lidocaine or corticosteroids.

To better select patients with facet joint pain, after patient recruitment, a test block of the medial branch of the facet joint was performed using 2% lidocaine, following the protocol of Rocha et al.[8].

After ten minutes, the patient was asked to assess the intensity of pain using the VAS. If the pain improved by 50% or more, the block was considered positive and the patient was definitively included in the study. On the other hand, if after ten minutes the evaluated patient reported an improvement of less than 50% in pain on the VAS, the patient was excluded from the study.

After the patient selection, they were randomized into two blocks: Intervention with PRP or Intervention with corticosteroid, and they were kept blinded, not knowing which treatment they had received, just like the researcher, who administered the questionnaires.

For the preparation of PRP, peripheral venous blood was collected under aseptic technique in tubes containing anticoagulant of the type Citric Acid, Sodium Citrate, and Dextrose (ACD). The preparation followed the protocol with double centrifugation, as described by Machado et al.[9].

The needles were positioned in the topography of the medial branches of the dorsal rami at L3-L4, L4-L5, and L5-S1 bilaterally, totaling six facet joints (Figure 1). Two mL of PRP or corticosteroid (Methylprednisolone Acetate) was applied to each articular facet. It is worth noting that, in both groups, one mL of 25% bupivacaine was also administered together with the injection in each articular facet.

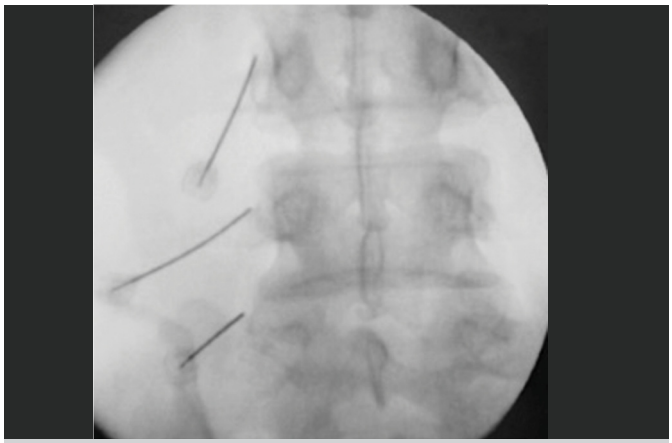


Figure 1. X-ray image of the procedure. X-ray image showing needle placement along the course of the medial branches of the dorsal rami at L3-L4, L4-L5, and L5-S1 on the left side.

Epidemiological data such as age, sex, BMI, physical activity, and presence of comorbidities were analyzed, in addition to clinical evaluations before the procedure, one month, and three months after the infiltration. During these periods, the Oswestry Disability Index (ODI)[10] and Roland Morris (RM)[11], as well as the Visual Analog Scale (VAS)[12], were evaluated.

The interaction effects between the evaluation moment and treatment, the intervention effect, and the time effect were assessed through a repeated measures analysis of variance (ANOVA). In all statistical analyses, a significance level of 5% was adopted, meaning that results with a p-value less than 5% ($p < 0.05$) were considered statistically significant.

The research project was submitted for evaluation to the Ethics Committee on the Brazil Platform under CAAE 76715617.0.2001.5479 and approved by opinion 4.033.153.

It is emphasized that the refusal to accept the Informed Consent Form (ICF) or withdrawal of consent can occur at any time, without any consequences of any kind to the individual. Finally, all institutional rules established by resolutions 466/2012 and 510/2016 were strictly respected.

RESULTS

Total Sample

A total of 92 patients were evaluated, and of these, those who presented positive block ($n = 60$) were included. However, one of these patients did not attend the follow-up for evaluation, and thus was excluded from the protocol.

According to Table 1, it can be observed that the two groups were homogeneous concerning the group in all evaluated variables ($p > 0.05$). It can also be observed that, regardless of the group, most patients were female with an average age of 56 years, and did not engage in physical activity. Furthermore, most patients had comorbidities, with hypertension being the most common.

Oswestry

Table 2 shows the descriptive measures of the Oswestry Disability Index (ODI) according to treatment and follow-up time. It can be observed that there was a significant interaction effect between treatment and time ($p = 0.038$), meaning that the two treatments did not exhibit the same behavior of the ODI score over time on average.

This behavior can be seen in Figure 2, where it is noted that from the first to the third month, there is a reversal in the behavior of the average ODI score ($p = 0.011$), in which in the first month the

Table 1. Distribution of demographic and clinical data of patients evaluated in the study.

Variables	Groups		p-value
	PRP (n=27)	Corticosteroid (n=32)	
Sex			**0.200**
Male	10 (37.0%)	7 (21.9%)	
Female	17 (63.0%)	25 (78.1%)	
Age, years			**0.907**
Mean ± standard deviation	55.8 10.7	56.1 10.0	
IMC, Kg/m²			**0.733**
Mean ± standard deviation	28.5 4.3	28.1 4.0	
Engages in physical activity			**0.517**
No	17 (63.0%)	22 (71.0%)	
Yes	10 (37.0%)	9 (29.0%)	
Comorbidities			**0.192**
Absent	12 (44.4%)	9 (28.1%)	
Present	15 (55.6%)	23 (71.9%)	
Hypertension	11 (40.7%)	13 (40.6%)	0.993
Diabetes	7 (25.9%)	6 (18.8%)	0.508
Hypercholesterolemia	7 (25.9%)	5 (15.6%)	0.327
Hypothyroidism	1 (3.7%)	0 (0.0%)	0.458
Hyperthyroidism	1 (3.7%)	1 (3.1%)	>0.999
Kidney failure	0 (0.0%)	0 (0.0%)	-
Asma/DPOC	1 (3.7%)	2 (6.3%)	>0.999
Depression	1 (3.7%)	1 (3.1%)	>0.999
Rheumatoid arthritis	2 (7.4%)	2 (6.3%)	>0.999
Fibromyalgia	2 (7.4%)	5 (15.6%)	0.437
Others	3 (11.2%)	8 (25.0%)	0.172
Levels with osteoarthritis			**0.190**
1	1 (3.7%)	6 (19.4%)	
2	9 (33.3%)	4 (12.9%)	
3	9 (33.3%)	13 (41.9%)	
4	5 (18.5%)	6 (19.4%)	
5	3 (11.2%)	2 (6.4%)	

Table 2. Mean ± standard deviation of the ODI score according to treatment and time.

Treatment	Time			p-value
	Baseline	1 month	3 months	
PRP	21.2 ± 7.3	22.0 ± 8.0	17.4 ± 9.7	0.016
Corticosteroid	21.0 ± 8.3	18.7 ± 8.0	20.0 ± 1.6	0.256
p-value	0.933	0.102	0.299	

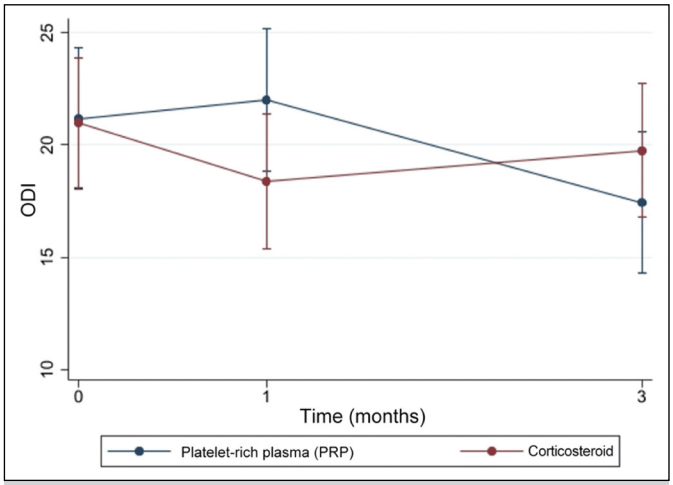


Figure 2. ODI Score. Mean profile (95% CI) of the ODI score over time according to treatment.

average observed in the PRP treatment was higher than that of the corticosteroid treatment, but in the third month the average observed in the PRP treatment was lower than that of the corticosteroid treatment.
On the other hand, it was found that there was no significant difference in the average ODI score between the two treatments at any of the evaluation moments ($p > 0.05$).

Visual Analog Scale

Table 3 shows the descriptive measures of the Visual Analog Scale (VAS) for pain according to treatment and follow-up time. It can be observed that there was no significant interaction effect between treatment and time ($p = 0.119$), meaning that the two treatments exhibited the same behavior of VAS over time on average. Furthermore, it was found that there was no significant difference between the two treatments over time ($p = 0.689$). Figure 3 illustrates the average profile of the treatments throughout the evaluation.

Table 3. Mean ± standard deviation of the visual analog scale (VAS) of pain according to treatment and time.

Treatment	Time			p-value
	Baseline	1 month	3 months	
PRP	71.0 ± 23.1	59.9 ± 25.3	52.1 ± 24.8	0.012
Corticosteroid	76.9 ± 23.2	51.1 ± 27.5	60.4 ± 29.4	<0.001
p-value	0.391	0.210	0.225	

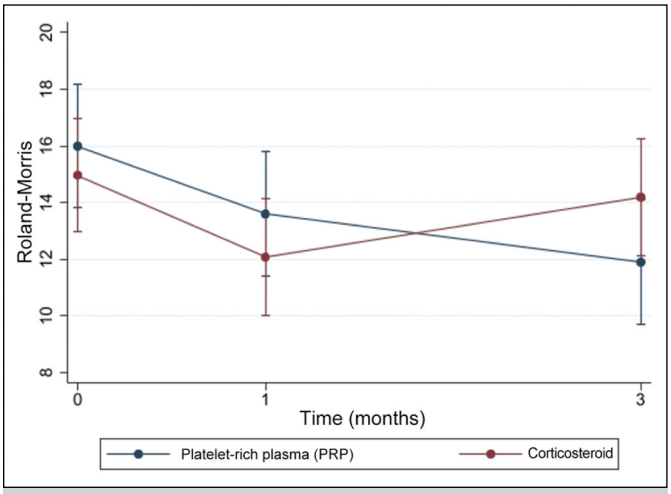


Figure 3. Visual Analog Scale for Pain. Mean profile (95% CI) of VAS over time according to treatment.

Roland-Morris

Table 4 shows the descriptive measures of the Roland-Morris score according to treatment and follow-up time. It can be observed that there was a significant interaction effect between treatment and time ($p = 0.010$), meaning that the two treatments did not exhibit the same behavior of the Roland-Morris score over time on average. This can be seen in Figure 4, where it is noted that from the first to the third month, there is a reversal in the behavior of the average Roland-Morris score ($p = 0.010$), meaning that in the first month the average observed of the Roland-Morris score in the PRP treatment was higher than that of the corticosteroid treatment, but in the third month the average observed in the PRP treatment was lower than that of the corticosteroid treatment.
However, it was found that there was no significant difference in the average Roland-Morris score between the two treatments at any of the evaluation moments ($p > 0.05$).

Table 4. Mean ± standard deviation of the Roland-Morris score according to treatment and time.

Treatment	Time			p-value
	Baseline	1 month	3 months	
PRP	16.0 ± 6.0	13.6 ± 5.8	11.9 ± 5.4	<0.001
Corticosteroid	15.0 ± 4.5	12.3 ± 5.6	14.4 ± 7.2	0.013
p-value	0.498	0.321	0.133	

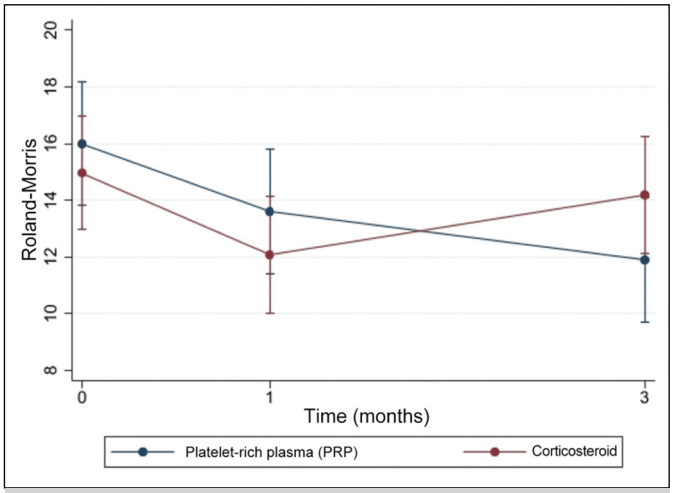


Figure 4. Roland-Morris score. Mean profile (95% CI) of the Roland-Morris score over time according to treatment.

DISCUSSION

The literature shows that facet joint disorders are responsible for 15% to 50% of cases of chronic low back pain[13,14]. In our study, of the 92 patients with chronic low back pain who underwent test infiltration of the medial branch of the facet joint, 60 (65%) patients reported pain relief (VAS ≥ 50%), suggesting pain of facet origin. In line with the literature, the study showed a prevalence of female sex in both studied groups, PRP and corticosteroid (63% and 78.1%, respectively), as well as an average age over 50 years[15]. It is worth noting the clinical homogeneity of the patients in all evaluated epidemiological aspects across the two groups, highlighting that the randomization was effective.

In the functional assessment, both the Oswestry score and the Roland-Morris score showed similar patterns in the present study. Although they did not show statistical significance in superiority of PRP compared to corticosteroids, both showed interaction between treatment and time; that is, PRP and corticosteroids exhibited different behaviors over time in the evaluation of these scores. It is important to remember that in these functional scores, the higher the score, the more debilitated the patient is.

When analyzing the graphs of the functional variables (ODI and RM) (Graphs 1 and 3), we noted that a factor that may have been crucial in the lack of observed superiority of the PRP group over the corticosteroid group was the absence of longer follow-up. Given that in both graphs, the trend was for the curves between the PRP group and the corticosteroid group to diverge, we can assume that if we had evaluated the patients at 6 months, some statistical difference might have been evidenced.

In a randomized clinical trial with 46 patients conducted by Wu et al.[16], which also compared PRP and corticosteroids in patients with facet joint pain, the same trend in ODI and RM functional scores was observed, but with statistical significance.

When we evaluated the pain scale (VAS), this trial did not show a significant difference between PRP and corticosteroids at any of the evaluated time points, nor did it present interaction between time and treatment. This finding does not confirm what other trials have already evidenced regarding the superiority of PRP over corticosteroids in pain relief, especially at 3 and 6 months[16,17].

Due to the scarcity of literature on PRP in facet syndrome, we greatly benefit from studies on osteoarthritis of peripheral joints, among which the knee is the target of the largest number of publications on PRP in joints[18]. In the review by Dong et al.[18], a benefit was evidenced in the use of PRP in knee osteoarthritis, mainly in improving pain and function in the short and medium term (1, 2, 3, 6, and 12 months). The same was not evidenced for the use of PRP in hip osteoarthritis, where the findings were still conflicting regarding pain and function improvement, which was also observed in our study[18].

On the other hand, in a large recent randomized clinical trial not included in Dong's systematic review et al.[18], PRP was compared with placebo in knee osteoarthritis[19] and contrary to the findings of the systematic review, there was no difference at 12 months for pain outcomes and cartilage characteristics between the two groups. In this aspect, the lack of effect of PRP on pain assessment was also found in our trial.

This effect of non-superiority of PRP for pain improvement may reside in the fact that in both cases, the chosen times for follow-up may not have been adequate. It is known that the effect of PRP generally has its best effect between 3 and 12 months[3,20], and both in the present trial, where patients were evaluated only for 3 months, and in the clinical trial comparing PRP and placebo in knee osteoarthritis, where patients were followed up only at 2 and 12 months, this optimal window was not respected.

Evaluating the limitations that this study presents, we can first cite the sample size. However, even with a reduced sample, this is the largest series in the literature studying PRP in lumbar facet syndrome. Another limitation in the present study is related to the short follow-up. Considering these points, it is of fundamental importance that new randomized controlled clinical trials, with a homogeneous population, reproducible methodology, longer follow-up, and especially larger sample sizes be conducted, so that we can draw better conclusions about the two treatment methods in lumbar facet syndrome.

CONCLUSION

The present study showed results that evidence the absence of superiority of PRP compared to corticosteroids in pain assessment and functional scores in the treatment of lumbar facet syndrome up to 3 months, making it impossible to see a difference between the methods in this sample.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. WZR: data acquisition and drafting of the work; RSSY: conception of the work and interpretation of the data; ESM: conception and design of the work; RGMM: data analysis and critical revision; AG: data analysis and critical revision; RM: conception of the work and final approval of the manuscript.

DATA AVAILABILITY DECLARATION

The underlying content of the research text is not contained in the manuscript and will be available upon request from reviewers.

REFERENCES

1. Manek NJ, MacGregor AJ. Epidemiology of back disorders: prevalence, risk factors, and prognosis. Curr Opin Rheumatol. 2005;17(2):134-40. doi: 10.1097/01.bor.0000154215.08986.06.
2. Hancock MJ, Maher CG, Latimer J, Spindler MF, McAuley JH, Laslett M, et al. Systematic review of tests to identify the disc, SIJ or facet joint as the source of low back pain. Eur Spine J. 2007;16(10):1539-50. doi: 10.1007/s00586-007-0391-1.
3. Desai MJ, Mansfield JT, Robinson DM, Miller BC, Borg-Stein J. Regenerative Medicine for Axial and Radicular Spine-Related Pain: A Narrative Review. Pain Pract. 2020;20(4):437-53. doi: 10.1111/papr.12868.
4. Nichols AW. Complications associated with the use of corticosteroids in the treatment of athletic injuries. Clin J Sport Med. 2005;15(5):370-5. doi: 10.1097/01.jsm.0000179233.17885.18.
5. Zhu Y, Yuan M, Meng HY, Wang AY, Guo QY, Wang Y, et al. Basic science and clinical application of platelet-rich plasma for cartilage defects and osteoarthritis: a review. Osteoarthritis Cartilage. 2013;21(11):1627-37. doi: 10.1016/j.joca.2013.07.017.
6. Sanapati J, Manchikanti L, Atluri S, Jordan S, Albers SL, Pappolla MA, et al. Do Regenerative Medicine Therapies Provide Long-Term Relief in Chronic Low Back Pain: A Systematic Review and Metaanalysis. Pain Physician. 2018;21(6):515-40.
7. Schneider BJ, Hunt C, Conger A, Qu W, Maus TP, Vorobeychik Y, et al. The effectiveness of intradiscal biologic treatments for discogenic low back pain: a systematic review. Spine J. 2022;22(2):226-37. doi: 10.1016/j.spinee.2021.07.015.
8. Rocha ID, Cristante AF, Marcon RM, Oliveira RP, Letaif OB, Barros Filho TE. Controlled medial branch anesthetic block in the diagnosis of chronic lumbar facet joint pain: the value of a three-month follow-up. Clinics (Sao Paulo). 2014;69(8):529-34. doi: 10.6061/clinics/2014(08)05.
9. Machado ES, Soares FP, Yamaguchi RS, Felipone WK, Meves R, Souza TAC, et al. A Simple Double-Spin Closed Method for Preparing Platelet-Rich Plasma. Cureus. 2022;14(1):e20899. doi: 10.7759/cureus.20899.
10. Vigatto R, Alexandre NM, Correa Filho HR. Development of a Brazilian Portuguese version of the Oswestry Disability Index: cross-cultural adaptation, reliability, and validity. Spine (Phila Pa 1976). 2007;32(4):481-6. doi: 10.1097/01.brs.0000255075.11496.47.
11. Júnior JJS, Nicholas MK, Pimenta CA de M, Asghari A, Thieme AL. Validação do Questionário de Incapacidade Roland Morris para dor em geral. Rev Dor. 2010;11(1):28-36.
12. Chiarotto A, Maxwell LJ, Ostelo RW, Boers M, Tugwell P, Terwee CB. Measurement Properties of Visual Analogue Scale, Numeric Rating Scale, and Pain Severity Subscale of the Brief Pain Inventory in Patients With Low Back Pain: A Systematic Review. J Pain. 2019;20(3):245-63. doi: 10.1016/j.jpain.2018.07.009.
13. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. The relative contributions of the disc and zygapophyseal joint in chronic low back pain. Spine (Phila Pa 1976). 1994;19(7):801-6. doi: 10.1097/00007632-199404000-00013.
14. Cohen SP, Raja SN. Pathogenesis, diagnosis, and treatment of lumbar zygapophysial (facet) joint pain. Anesthesiology. 2007;106(3):591-614. doi: 10.1097/00000542-200703000-00024.
15. Manchikanti L, Boswell MV, Singh V, Pampati V, Damron KS, Beyer CD. Prevalence of facet joint pain in chronic spinal pain of cervical, thoracic, and lumbar regions. BMC Musculoskelet Disord. 2004;5:15. doi: 10.1186/1471-2474-5-15.
16. Wu J, Zhou J, Liu C, Zhang J, Xiong W, Lv Y, et al. A Prospective Study Comparing Platelet-Rich Plasma and Local Anesthetic (LA)/Corticosteroid in Intra-Articular Injection for the Treatment of Lumbar Facet Joint Syndrome. Pain Pract. 2017;17(7):914-24. doi: 10.1111/papr.12544.
17. Byvaltsev VA, Kalinin AA, Okoneshnikova AK, Satardinova EE. Analiz klinicheskoĭ éffektivnosti primeneniia PRP-terapii pri lechenii patsientov s izolirovannym fasetochnym sindromom poiasnichnogo otdela pozvonochnika [Analysis of the clinical efficacy of platelet-rich plasma therapy in the treatment of patients with isolated facet-syndrome of the lumbar spine]. Zh Nevrol Psikhiatr Im S S Korsakova. 2019;119(11):27-31. Russian. doi: 10.17116/jnevro201911911127.
18. Dong Y, Zhang B, Yang Q, Zhu J, Sun X. The effects of platelet-rich plasma injection in knee and hip osteoarthritis: a meta-analysis of randomized controlled trials. Clin Rheumatol. 2021;40(1):263-77. doi: 10.1007/s10067-020-05185-2.
19. Bennell KL, Paterson KL, Metcalf BR, Duong V, Eyles J, Kasza J, et al. Effect of Intra-articular Platelet-Rich Plasma vs Placebo Injection on Pain and Medial Tibial Cartilage Volume in Patients With Knee Osteoarthritis: The RESTORE Randomized Clinical Trial. JAMA. 2021;326(20):2021-30. doi: 10.1001/jama.2021.19415.
20. Filardo G, Previtali D, Napoli F, Candrian C, Zaffagnini S, Grassi A. PRP Injections for the Treatment of Knee Osteoarthritis: A Meta-Analysis of Randomized Controlled Trials. Cartilage. 2021;13(1_suppl):364S-75S. doi: 10.1177/1947603520931170.

DOI: http://dx.doi.org/10.1590/1413-785220263404e301170

VALIDATION OF AN OPTICAL METHOD FOR MEASURING THE GAP BETWEEN SUTURED TENDON STUMPS

VALIDAÇÃO DE UM MÉTODO ÓPTICO PARA MENSURAÇÃO DO ESPAÇO ENTRE COTOS TENDÍNEOS SUTURADOS

Luiz Sorrenti[1], *Elcio Koodiro Yoshida*[1], *Antonio Isidoro de Sousa Neto*[1], *César Augusto Martins Pereira*[1], *Hugo Alberto Nakamoto*[1], *Rames Mattar Junior*[1]

1. Universidade de Sao Paulo, Faculdade de Medicina, Hospital das Clinicas, (HC-FMUSP), Instituto de Ortopedia e Traumatologia, Hand Surgery and Reconstructive Microsurgery Group, Sao Paulo, SP, Brazil.

ABSTRACT

Objective: This study aims to investigate the accuracy and precision of an optical measurement method based on backlight video analysis for measuring gap distances between tendon stumps in a rigid tendon model. Methods: A rigid 3D-printed tendon model was sutured and subjected to traction using a universal mechanical testing machine. The gap between stumps was filmed in frontal and lateral planes with backlight illumination and a simple mirror system. A custom software program calculated gap distances from the recorded videos and the optical measurements were compared with mechanical testing machine measurements to assess precision and accuracy. Results: The mean difference (bias) in the frontal plane was -0.01 mm (SD: 0.01 mm) and in the lateral plane was 0.04 mm (SD: 0.03 mm). Bland-Altman analysis showed limits of agreement (LoA) of 0.04 mm in the frontal plane and 0.06 mm in the lateral plane. Measurements between planes showed a very strong correlation with a Spearman's rho correlation test of 1.000 ($p<0.001$). Conclusion: The optical method showed errors on the order of hundredths of a millimeter and proved to be both precise and accurate, validating a new method for gap measurement in biomechanical tendon studies. ***Level of Evidence IV; Experimental Validation Study.***

Keywords: Tendons; Tendon Transfer; Biomechanical Phenomena; Suture Techniques.

RESUMO

Objetivo: Este estudo tem como objetivo investigar a acurácia e a precisão de um método de mensuração óptica baseado na análise de vídeo com retroiluminação para medir as distâncias do afastamento entre cotos tendíneos em um modelo tendíneo rígido. Métodos: Um modelo tendíneo rígido impresso em 3D foi suturado e submetido à tração em uma máquina universal de ensaios mecânicos. O espaço entre os cotos foi filmado nos planos frontal e lateral com retroiluminação e um sistema simples de espelho. Um software personalizado calculou as distâncias do espaço a partir dos vídeos gravados, e as mensurações ópticas foram comparadas com as mensurações da máquina de ensaios mecânicos para avaliar a precisão e a acurácia. Resultados: A diferença média, ou viés, no plano frontal foi de -0,01 mm (DP: 0,01 mm) e, no plano lateral, de 0,04 mm (DP: 0,03 mm). A análise de Bland-Altman apresentou limites de concordância (LoA) de 0,04 mm no plano frontal e de 0,06 mm no plano lateral. As mensurações entre os planos apresentaram correlação muito forte, com coeficiente rho de Spearman de 1,000 ($p<0,001$). Conclusão: O método óptico apresentou erros da ordem de centésimos de milímetro e demonstrou ser preciso e acurado, validando um novo método para mensuração do espaço em estudos biomecânicos de tendões. ***Nível de Evidência IV; Estudo de Validação Experimental.***

Descritores: *Tendões; Transferência Tendinosa; Fenômenos Biomecânicos; Técnicas de Sutura.*

Citation: Sorrenti L, Yoshida EK, Sousa Neto AI, Pereira CAM, Nakamoto HA, Mattar Junior R. Validation of an optical method for measuring the gap between sutured tendon stumps. Acta Ortop Bras. [online]. 2026;34(4): Page 1 of 6. Available from URL: http://www.scielo.br/aob.

INTRODUCTION

Tendon is a tissue formed mainly by collagen (65-80%), elastin (1-2%) produced by tenocytes, surrounded by a proteoglycan matrix responsible for transmitting muscle force to joints and bones.[1] It has the capacity to adapt to mechanical loads by modifying its structure, composition and mechanical properties[2,3], however, its overload can result in acute or chronic injuries.[4] Injuries of flexor tendons of the fingers often result in functional disability[5] and it is known that resistant suturing techniques to allow early active mobilization are essential for healing, prevention of adhesion formation and better results.[6] Despite advances in suturing techniques and post-operative rehabilitation, an important complication is the formation of a tendon gap, the distance between the edges of the two tendon stumps

All authors declare no potential conflict of interest related to this article.

The study was conducted at the Universidade de Sao Paulo, Faculdade de Medicina, Hospital das Clinicas, (HC-FMUSP), Instituto de Ortopedia e Traumatologia, Hand Surgery and Reconstructive Microsurgery Group, Sao Paulo, SP, Brazil.
Correspondence: Luiz Sorrenti. 333, Rua Dr. Ovídio Pires de Campos, Sao Paulo, SP, Brazil. 05402-000. luiz.sorrenti@hc.fm.usp.br

Article received on 09/22/2025 approved on 02/18/2026.
Handling Editor: Flavio Faloppa

that are joined by the surgical suture, when subjected to traction forces, which can lead to poor functional results after tendon repair.[7] A gap greater than 2 mm after repair of the flexor digitorum profundus tendon significantly increases the resistance to sliding, leading to a higher risk of adhesion and the need for a future tenolysis.[6–8] Linnanmäki et al.[9] concluded that, when the suture is subjected to repeated stress, there is a suture fatigue point coinciding with the beginning of tendon plastic deformation that predicts the formation of gaps and then increases the risk of failure if repeated stress is maintained. Furthermore Gelberman et al.[10] established that the formation of a 3 mm gap is detrimental to the gain in strength and stiffness in tendon healing, using a canine model.

The formation of a gap in the tendon suture compromises the quality of the tenorrhaphy and we have observed the absence of new methods for measuring the formation of this gap.[7,11] Therefore, there is a need to update the methods for measuring the gap between tendon stumps in experimental tests, seeking greater precision of the measured distance between tendon stumps. Greater precision in this analysis can improve the quality of existing techniques for tendon suturing. It is expected that there will be a direct benefit to patients, through the standardization of the recording and evaluation of the gap in trials, research and development of new surgical techniques. The aim of this study is to validate the accuracy of the optical method for measurement of the distance, in the frontal and lateral planes, between the stumps of a 3D-printed tendon model in biomechanical experiments with the universal mechanical testing machine.

METHODS

Model

The study obtained gap measurements of the tendon stump model in the frontal and lateral planes by filming it with the help of a mirror reflecting the lateral face and backlighting the model in both planes. The Kratos K5002 universal mechanical testing machine (Kratos Equipamentos Industriais Ltda., São Paulo, São Paulo, Brazil) was used to precisely separate the tendon stump and acquire the distance of the movement. A Canon EOS Rebel T2i digital camera with a Canon EF-S 18-135mm f/3.5-5.6 IS STM lens (Canon Inc., Japan) recorded the footage. The images were processed using specific software developed for the study by the authors. The tendon stump model was made with the shape of an elliptical cylinder and split orthogonally to reproduce the faces of the stumps of a ruptured tendon (Figure 1). It was printed using a 3D Machine One 3D filament printer in white polylactic acid (PLA) plastic. The rigid plastic material was chosen so that there would be no significant deformation of the material during traction of the stumps by the universal mechanical testing machine.

The model was based on the thickness of the flexor digitorum profundus tendon of the adult human hand and has an elliptical cross-section with a major width of 7.9 mm and a minor width of 6.2 mm and a length (longitudinal axis) of 30 mm for each part of the stump.[12] Eight holes were made on the split sides of the stumps to allow the surgical sutures to pass through, simulating the tenorrhaphy of the tendon stumps repair (Figure 1).

The two parts of the tendon model were fixed to the universal mechanical testing machine by rigid clamps ensuring that the axis of the displacement aligned with the longitudinal axis of the stumps and no other movement could occur. The machine recorded a precise distance and time during the tests (Figure 2A, B and C). The two parts of the stump were connected simulating a four-strand tenorrhaphy using black nylon suture No. 4-0. The suture was tensioned with an elastic band and attached to the base of the model without causing excessive traction on the stump during its separation (Figure 2I).

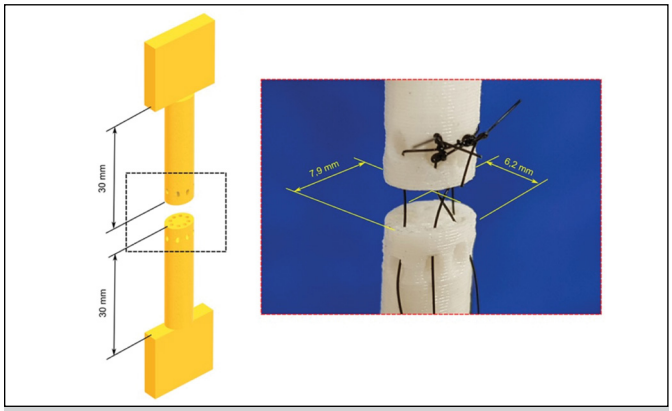


Figure 1. 3D-printed tendon model - Stump faces have eight holes for the tendon suture.

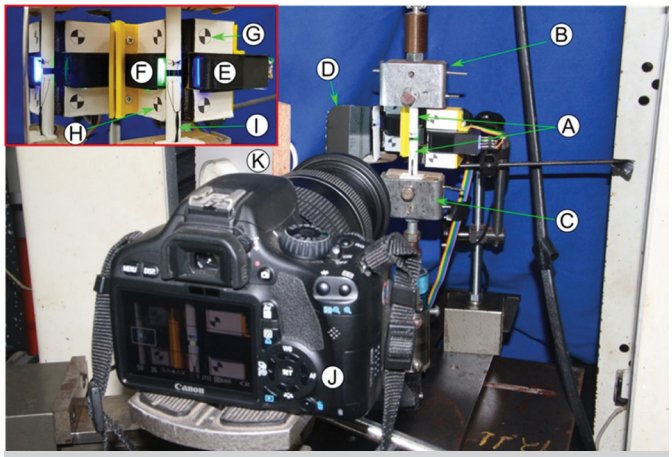


Figure 2. Assembly of the measuring system - Tendon model (A), fixation claws (B and C), metallic mirror (D), backlighting light sources (E and F), templates (G and H), camera (J) and light source (K). The sutures were tensioned with an elastic band (I) attached to the lower clamp (C).

To start the test, the faces of the model's stumps were brought together, setting the initial gap at zero. To record the gap optically in the frontal and lateral planes of the tendon model, the camera was positioned on a tripod at the same height as the model (Figure 2J). The framing of the footage was in landscape, with the tendon model centered, obtaining the frontal plane of the stumps, and on the left side the mirror image of the lateral face of the model, obtaining the lateral plane of the stumps (Figure 2D). An Indusbello IBM-001 metallic mirror (Indusbello, Londrina, Paraná, Brazil) was used so that the image would not be double reflected like a regular glass mirror. The gap measurements were obtained by using printed templates of known distances included in the same planes as the filmed image of interest (Figure 2G, H). The distances between the centers of the templates were previously obtained by the Deltronic DV 114 profile projector, with a linear accuracy of 0.005mm (Deltronic, Santa Ana, California, EUA), which allows the filming to be calibrated (Figure 3). Two LED projectors were positioned orthogonally to backlight the model to increase the luminous contrast with the passage of light through the gap formed when the stumps were separated. The front backlight of the stump was green and the mirrored side backlight was blue.

To ensure that the displacement and time data from the universal mechanical testing machine was synchronized with the filming, an LED was included which turned on when the displacement was made by the machine, marking the start of the movement when it turned on and the end when it turned off (Figure 3H).

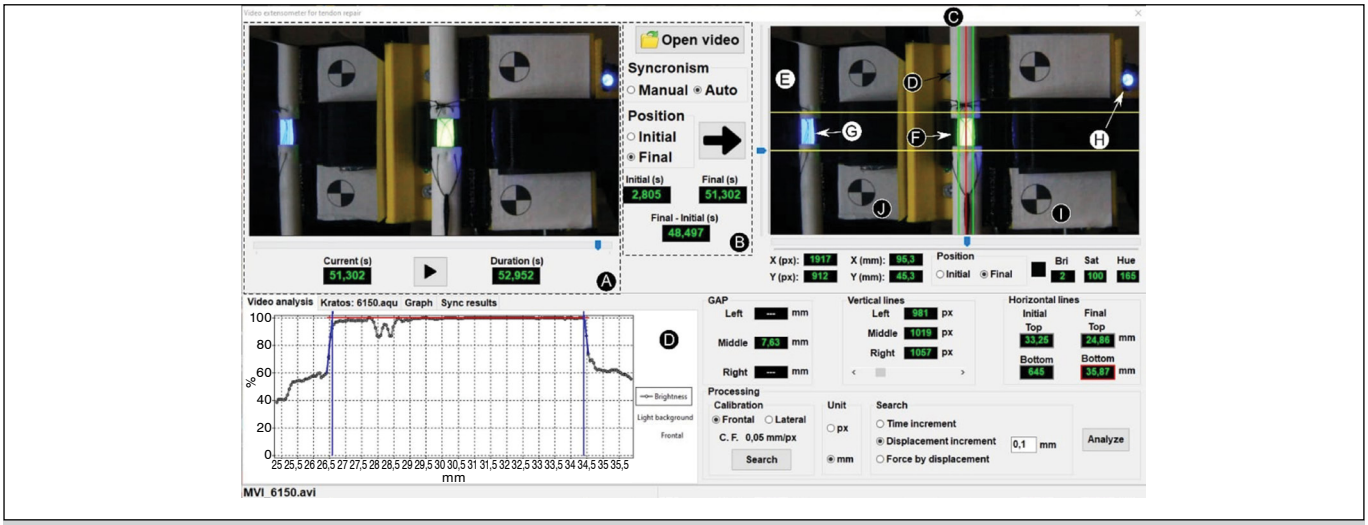


Figure 3. Measurement software - Video controls (A), position and synchronization controls (B). On the right (C) the tendon model video (D) and the lateral mirrored image of the tendon model (E), the synchronism LED (H), frontal (I) and lateral (J) template, frontal (F) and lateral (G) backlight source. The graph (D) represents the brightness (y-axis, %) corresponding to the red vertical line bounded by the two yellow horizontal lines (x-axis, mm).

Computer software based on the Free Pascal language was developed by the author within the Typhon integrated development environment (Video 1S - https://youtu.be/wAKF_sRU-pE). This software identifies the gap as the region between the stumps backlit by the projector during filming. The software used the fixed templates to determine the gap in millimeters. Each test was carried out at a speed of 10 mm/min until a total movement of 6 mm was recorded by the universal mechanical testing machine, with a displacement accuracy of hundredths of a millimeter. Ten tests were repeated with the tendon model. The displacement information measured by the machine was recorded using the Lynx ADS 2000 data acquisition system with a displacement resolution of 0.01 mm (Lynx Tecnologia Eletrônica Ltda., São Paulo, São Paulo, Brazil).

In order for the software to automatically take measurements of the gap in the tendon model, the user must set up the following within the program itself: synchronism, selection of the analysis area (position), calibration and selection of the search criteria for the gap.

Software settings:

Synchronism: Time synchronization of the footage with the data obtained by the universal mechanical testing machine. In the auto option, the user enters the approximate location of the synchronism LED in the image to recognize the start and end of the movement.

Selection of the analysis area (position): delimitation of the image for the evaluation and analysis of the gap formed, by positioning the green vertical lines and the yellow horizontal lines.

Calibration: In the calibration process, the software asks the user for the approximate location of the center of each template seen in the footage, and the centers of the templates are automatically identified. Calibration is defined as the ratio of the known distance of the template to the distance in pixels of the points found.

Gap search criteria: The software determines the search criteria, which can be by time increment; displacement increment according to the progression of the universal mechanical testing machine, a criteria used in the validation of the model; and by predefined gap (Force by GAP) with the recognition of the increase in the gap in the filming, interesting in comparative studies.

Methods of Assessment

The gap measurement carried out by the software is based on the sudden change in brightness caused by the contrast between the stump model and the background at the end of the stumps (Figure 4),

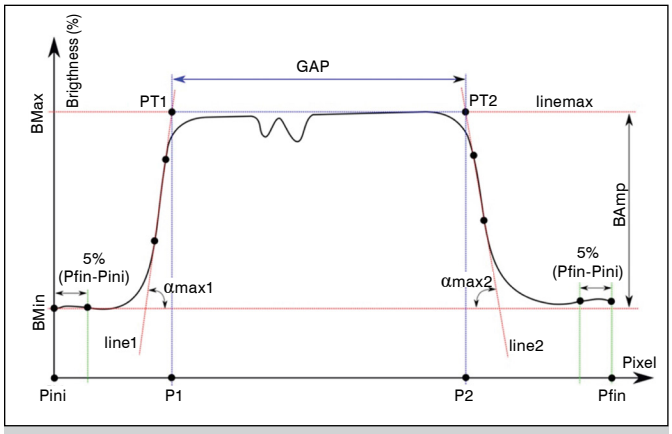


Figure 4. Graph representing the gap distance of the backlit tendon model - It shows the distribution of brightness measured vertically on the image between the initial (Pini) and final (Pfin) points.

and this interval is transformed into millimeters through the software's calibration with the fixed template distances included. The software calculates the average gap in the interval bounded by the green vertical and yellow horizontal lines, intervals manually selected before the start of the filming process (Figure 3), for both frontal and lateral analysis, ensuring that the gap region is contained.

Statistical Analysis

This study compared measurements obtained from footage processing against the measurements recorded by a universal mechanical testing machine, utilizing the Lynx ADS 2000 data acquisition system during the stumps separation. Descriptive statistics, including the mean difference (bias), median, standard deviation (SD), interquartile range (IQR), and standard error of the mean (SEM), were calculated from the differences between footage measurements and data acquisition system measurements. These statistics were reported for each of the ten individual runs and for the aggregate 600 data pairs in both the frontal and the lateral planes, with no missing values. A significance level (α) of 0.05 and a statistical power ($1-\beta$) of 0.80 were adopted for this study. The Bland-Altman analysis[13] was the primary method for assessing limits of agreement (LoA). These graphical and statistical plots

differences against averages to identify systematic bias and calculates the 95% LoA, within which 95% of differences are expected to fall. The adequacy of the sample size (600 pairs) for the Bland-Altman analysis was confirmed using the method proposed by Lu et al.[14] This retrospective assessment determined the maximum allowed difference with the given sample size. Spearman's rho correlation test was performed to assess the relationship between the frontal and lateral plane measurements.

Prior to assessing the agreement between the two measurement methods, the normality of the differences was evaluated using the Shapiro-Wilk test. For both frontal and lateral plane differences, the p-value was less than 0.05, indicating a non-normal distribution of the data (Figure 5). Although Bland-Altman analysis usually uses normally distributed data, it was employed in this study to determine the LoA to provide a descriptive measure of agreement, with the understanding that the interpretation of these limits acknowledges the non-normal distribution of the differences. Given the minimal variation observed between the methods, the standard deviation (SD), interquartile range (IQR), and standard error of the mean (SEM) were reported with three decimal places, reflecting a precision less than the data acquisition system's limit.

The methodology proposed by Lu et al.[14] was utilized to determine the maximum allowable difference in a scenario of 600 tests. This approach is usually used to evaluate the adequacy of the Bland-Altman analysis sample size by pre-determining alpha, beta, and a maximum acceptable difference. In this study, a retrograde analysis was performed to assess these parameters of a maximum acceptable difference for the aggregate test.

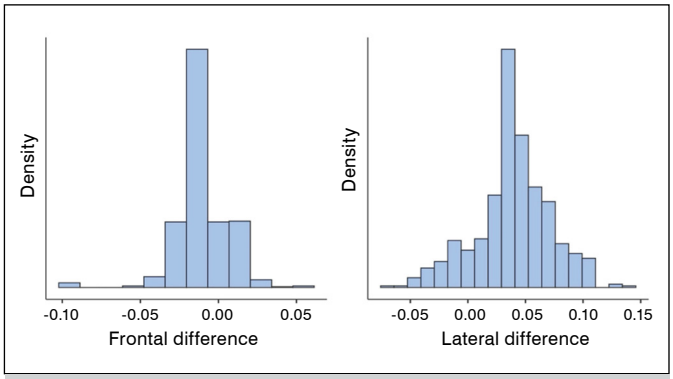


Figure 5. Frontal and lateral differences distribution - Graphics of differences of the measured gap and the displacement by the Kratos K5002 universal mechanical testing machine. With Shapiro-Wilk test with $p < 0.05$. Frontal plane on the left and lateral plane on the right.

RESULTS

Ten runs were performed with 60 measurements and a total of 600 pairs of measurement were recorded. The mean difference or bias of all data in the frontal plane was -0.01 mm (SD: 0.01 mm) and in the lateral plane was 0.04 mm (SD: 0.03 mm) (Table 1). It can be noted that the difference between the measurement performed by the method and the measurement obtained by the universal mechanical testing machine was on the order of hundredths of a millimeter in the frontal and lateral planes.

To describe the agreement between the collected data and universal mechanical testing machine, Bland-Altman analysis was performed using the Jamovi Version 2.6 statistical software (Jamovi, Sydney, Australia) with blandr module. The Bland-Altman analysis gave a frontal plane LoA of 0.04, upper LoA of 0.02 mm (Table 2, Figure 6) and lower LoA of -0.05 mm and a lateral plane LoA of 0.06, upper LoA of 0.10 mm and lower LoA of -0.02 mm (Table 3, Figure 7), and it was considered precise by the authors. With 600 tests and

Table 1. Descriptive data of differences of the measured gap and the displacement by the Kratos K5002 universal mechanical testing machine.

	Frontal Plane		Lateral Plane	
Run	Difference Mean SD; Difference Median (IQR)	SEM	Difference Mean SD; Difference Median (IQR)	SEM
1	-0.03 ± 0.012; -0.03 (0.010)	0.002	0.03 ± 0.031; 0.04 (0.030)	0.004
2	-0.02 ± 0.014; -0.02 (0.010)	0.002	0.04 ± 0.032; 0.04 (0.030)	0.004
3	-0.01 ± 0.011; -0.01 (0.010)	0.001	0.04 ± 0.027; 0.04 (0.033)	0.004
4	-0.01 ± 0.015; -0.01 (0.000)	0.002	0.04 ± 0.032; 0.04 (0.033)	0.004
5	-0.01 ± 0.015; -0.01 (0.010)	0.002	0.04 ± 0.031; 0.05 (0.040)	0.004
6	-0.01 ± 0.009; -0.01 (0.000)	0.001	0.03 ± 0.032; 0.04 (0.040)	0.004
7	-0.03 ± 0.013; -0.03 (0.010)	0.002	0.03 ± 0.030; 0.03 (0.043)	0.004
8	0 ± 0.012; 0 (0.010)	0.002	0.05 ± 0.029; 0.06 (0.033)	0.004
9	0.02 ± 0.010; 0.02 (0.010)	0.001	0.06 ± 0.031; 0.06 (0.042)	0.004
10	-0.02 ± 0.013; -0.01 (0.010)	0.002	0.04 ± 0.038; 0.04 (0.042)	0.005
All	-0.01 ± 0.018; -0.01 (0.020)	0.001	0.04 ± 0.033; 0.04 (0.040)	0.001

Standard deviation (SD); interquartile (IQR); Standard Error Mean (SEM).

Table 2. Bland-Altman analysis of frontal plane measurements with 95% confidence interval of limit of agreement.

		95% Confidence Interval	
	Estimate	Lower	Upper
Bias (n = 600)	-0.011	-0.012	-0.010
Lower limit of agreement	-0.046	-0.049	-0.044
Upper limit of agreement	0.024	0.022	0.027

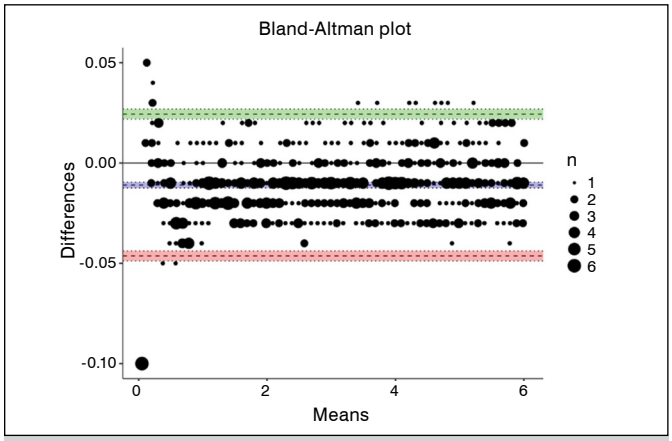


Figure 6. Bland-Altman plot of frontal plane measurements - Bland-Altman plot with upper limit of agreement (LoA) and confidence intervals of 95% in green, lower LoA and confidence intervals in red and bias and confidence intervals of 95% in blue. Circle size represent the density.

predetermined alpha and beta, the Lu et al. analysis showed a maximum allowed difference (Δ) of 0.05 mm for the frontal plane and a value of 0.11 mm for the lateral plane. The Med Calc® software (MedCalc Software Ltd, Ostend, Belgium) was used to perform this test. The Bland-Altman analysis was in agreement with the Lu et al. test[14], with lower values.

These data show that the average gap measurement in the lateral reflected plane was not as accurate and precise as that in the frontal plane, with an average difference of 0.05 mm and with the standard deviation being small as 0.03 mm between the planes. The Spearman rho's correlation test was very strong and with a significant correlation (1.000; $p<0.001$). No major complications or unexpected failures occurred during testing.

Table 3. Bland-Altman analysis of lateral plane measurements with 95% confidence interval of limit of agreement.

		95% Confidence Interval	
	Estimate	Lower	Upper
Bias (n = 600)	0.040	0.037	0.042
Lower limit of agreement	-0.025	-0.029	-0.020
Upper limit of agreement	0.104	0.099	0.108

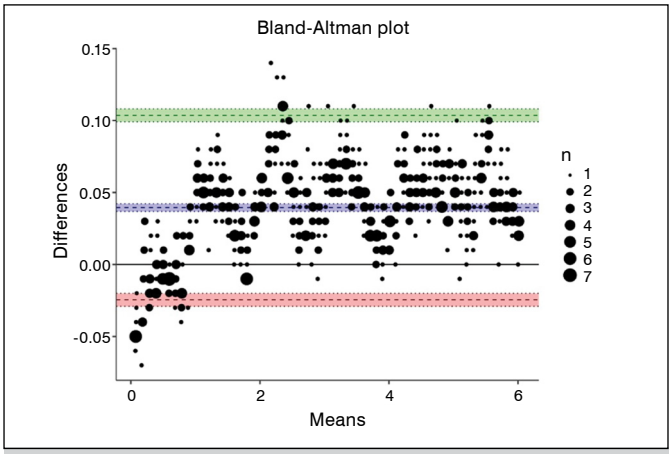


Figure 7. Bland-Altman plot of lateral plane measurements - Bland-Altman plot with upper limit of agreement (LoA) and confidence intervals of 95% in green, lower LoA and confidence intervals in red and bias and confidence intervals of 95% in blue. Circle size represent the density.

DISCUSSION

The determination of the precise distance between tendon stumps during biomechanical strength testing is essential for the development of a surgical technique that allows early active movement in the rehabilitation of tendon injuries, especially in the flexors of the fingers [15] But several mechanisms have been described in the literature for measuring the distance formed between tendon stumps in comparative tests of different tendon suturing techniques and most are imprecise, on the order of a millimeter, or very expensive and complex.[6,12]

Zhao et al.[7] and Lee et al.[16] recorded the formation of the gap using a digital video camera at a speed of 30 frames per second, associated with a micrometer ruler placed parallel and coplanar to the tendon. Croog et al.[17] used digital images at 5 frames per second, filmed using a ruler along the tendon. Lawrence et al.[18] used visual inspection with the aid of magnifying glasses with 2 to 3 times magnification to determine the moment when the gap occurs. Wong et al.[19] use photographic images associated with a ruler. In all cases, the measurements can suffer great inter-observer and even intra-observer variation.

In the case of Linnanmäki et al.[9], the gaps were measured statically after load application using camera photos, in comparison with photos taken previously, using computer software, but without the use of backlight for measurement, which does not prevent measurement errors due to changes in brightness and shadows. Yang et al.[20] compared the force required to generate a 2 mm gap in different types of sutures and used a caliper to perform the measurements, which can lead to variations and inaccuracies in the results.

No methods found in the literature associate the measurement of the gap in two orthogonal planes at the same time, frontal and lateral, through the analysis of images of the tendon backlighted with a light source. These methods use image analysis with the tendon illuminated in a single plane (frontal) and require a more sophisticated image analysis to detect the gap, since factors such as lighting, color of the object and of the background can interfere in the process.

In addition, there is an established correlation between the number of strands and the tension force of a tendon repair.[19] Therefore, sutures with different numbers of strands should be compared. In measurements taken with tenorrhaphy sutures, it was noted that they did not interfere with the gap measurements. The new method allows comparisons to be made between sutures with different numbers of passes, without interfering with the gap measurements. It is also possible in tests involving suture resistance, to correlate the size of the gap formed and the force required for it with the universal mechanical testing machine. Validation of the measurement system using a rigid plastic model and rigid clamps without resistance was necessary because the use of natural tendon would not allow direct correspondence between the data from the displacement meter of the universal mechanical testing machine and those recorded by the camera, due to their intrinsic elasticity and deformation before separation occurs. Also, for this reason, an isolated analysis of the measurement of the separation of the machine's data is not useful in biomechanical tests with natural tendons subjected to repair. These validations used a rigid model and future extrapolation to biological tissues should be cautious, considering tissue elasticity and deformation.

Validation of the method allows for full reliability in the measurements provided by the images and analyzed by the software, creating an accurate measurement method on the order of hundredths of a millimeter. The validation method is reproducible, low-cost, and capable of simultaneous dual-plane evaluation.

CONCLUSION

The optical method offers a reliable and accurate tool for measuring gap distances between tendon stumps on the order of hundredths of a millimeter. It is a promising technique for biomechanical research involving natural tendon repair evaluation.

ACKNOWLEDGEMENTS

We would like to thank Luciano Ruiz Torres and Erick Yoshio Wataya for their assistance and guidance in this research.

CONTRIBUTIONS OF THE AUTHORS

Each author contributed individually and significantly to the development of this article. LS: writing – original draft, writing – review and editing, conceptualization, methodology, investigation, data curation, visualization, formal analysis, project administration, resources, and software. EKY: writing – original draft, conceptualization, methodology, investigation, visualization, formal analysis, and resources. AISN: writing – original draft, conceptualization, methodology, investigation, visualization, and resources. CAMP: writing – original draft, conceptualization, methodology, investigation, visualization, resources, and software. HAN: writing – review and editing, conceptualization, methodology, and investigation. RMJ: supervision, writing – original draft, writing – review and editing, conceptualization, methodology, investigation, data curation, formal analysis, project administration, resources, and software.

DATA AVAILABILITY DECLARATION

The authors confirm that the data supporting the findings of this study are available within the article. In addition, the datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

1. Kannus P. Structure of the tendon connective tissue. Scand J Med Sci Sports. 2000;10(6):312-20. doi: 10.1034/j.1600-0838.2000.010006312.x.
2. Wang JH. Mechanobiology of tendon. J Biomech. 2006;39(9):1563-82. doi: 10.1016/j.jbiomech.2005.05.011.
3. Wang JH, Guo Q, Li B. Tendon biomechanics and mechanobiology--a minireview of basic concepts and recent advancements. J Hand Ther. 2012;25(2):133-40. doi: 10.1016/j.jht.2011.07.004.
4. Kannus P, Paavola M, Paakkala T, Parkkari J, Järvinen T, Järvinen M. Pathophysiology of overuse tendon injury. Radiologe. 2002;42(10):766-70. doi: 10.1007/s00117-002-0795-1.
5. Gundlach BK, Zelouf DS. Flexor tendon reconstruction. Hand Clin. 2023;39(2):193-201. doi: 10.1016/j.hcl.2022.08.020.
6. Tang JB, Lalonde D, Harhaus L, Sadek AF, Moriya K, Pan ZJ. Flexor tendon repair: recent changes and current methods. J Hand Surg Eur Vol. 2022;47(1):31-9. doi: 10.1177/17531934211053757.
7. Zhao C, Amadio PC, Tanaka T, Kutsumi K, Tsubone T, Zobitz ME, et al. Effect of gap size on gliding resistance after flexor tendon repair. J Bone Joint Surg Am. 2004;86(11):2482-8. doi: 10.2106/00004623-200411000-00019.
8. Angeles JG, Heminger H, Mass DP. Comparative biomechanical performances of 4-strand core suture repairs for zone II flexor tendon lacerations. J Hand Surg Am. 2002;27(3):508-17. doi: 10.1053/jhsu.2002.32619.
9. Linnanmäki L, Göransson H, Havulinna J, Sippola P, Karjalainen T, Leppänen OV. Gap formation during cyclic testing of flexor tendon repair. J Hand Surg Am. 2018;43(6):570.e1-570.e8. doi: 10.1016/j.jhsa.2017.12.005.
10. Gelberman RH, Boyer MI, Brodt MD, Winters SC, Silva MJ. The effect of gap formation at the repair site on the strength and excursion of intrasynovial flexor tendons. An experimental study on the early stages of tendon-healing in dogs. J Bone Joint Surg Am. 1999;81(7):975-82. doi: 10.2106/00004623-199907000-00010.
11. Haddad R, Peltz TS, Lau A, Bertollo N, Nicklin S, Walsh WR. The relationship between gap formation and grip-to-grip displacement during cyclic testing of repaired flexor tendons. J Biomech. 2010;43(14):2835-8. doi: 10.1016/j.jbiomech.2010.06.006.
12. Mao WF, Wu YF, Zhou YL, Tang JB. A study of the anatomy and repair strengths of porcine flexor and extensor tendons: are they appropriate experimental models?. J Hand Surg Eur Vol. 2011;36(8):663-9. doi: 10.1177/1753193411414117.
13. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies J R Stat Soc Ser D (The Statistician). 1983;32(3):307-17. doi: 10.2307/2987937.
14. Lu MJ, Zhong WH, Liu YX, Miao HZ, Li YC, Ji MH. Sample size for assessing agreement between two methods of measurement by Bland-Altman method. Int J Biostat. 2016;12(2):20150039. doi: 10.1515/ijb-2015-0039.
15. Pan ZJ, Pan L, Xu YF, Ma T, Yao LH. Outcomes of 200 digital flexor tendon repairs using updated protocols and 30 repairs using an old protocol: experience over 7 years. J Hand Surg Eur Vol. 2020;45(1):56-63. doi: 10.1177/1753193419883579.
16. Lee HI, Lee JS, Kim TH, Chang SH, Park MJ, Lee GJ. Comparison of flexor tendon suture techniques including 1 using 10 strands. J Hand Surg Am. 2015;40(7):1369-76. doi: 10.1016/j.jhsa.2015.03.021.
17. Croog A, Goldstein R, Nasser P, Lee SK. Comparative biomechanic performances of locked cruciate four-strand flexor tendon repairs in an ex vivo porcine model. J Hand Surg Am. 2007;32(2):225-32. doi: 10.1016/j.jhsa.2006.11.009.
18. Lawrence TM, Davis TR. A biomechanical analysis of suture materials and their influence on a four-strand flexor tendon repair. J Hand Surg Am. 2005;30(4):836-41. doi: 10.1016/j.jhsa.2005.03.011.
19. Wong YR, Lee CS, Loke AM, Liu X, Suzana MJ I, Tay SC. Comparison of flexor tendon repair between 6-strand Lim-Tsai with 4-strand cruciate and Becker technique. J Hand Surg Am. 2015;40(9):1806-11. doi: 10.1016/j.jhsa.2015.05.007.
20. Yang W, Li J, Su Y, Liang W, Ren Y, Dong Y, et al. A modified flexor tendon suture technique combining kessler and loop lock flexor tendon sutures. Clinics (Sao Paulo). 2021;76:e2358. doi: 10.6061/clinics/2021/e2358.