Optimizing care in lumbar radiculopathy and neurogenic claudication: from injection to inference, and from clinician to algorithm

Verheijen, E.J.A.

Optimizing Care in Lumbar Radiculopathy and Neurogenic Claudication

From injection to inference, and from clinician to algorithm

Eduard J.A. Verheijen

Eduard J.A. Verheijen

Optimizing Care in Lumbar Radiculopathy and Neurogenic Claudication.
From injection to inference, and from clinician to algorithm

PhD thesis, Leiden University, Leiden, The Netherlands

Printing:	ProefschriftMaken | proefschriftmaken.nl
Layout and design:	ProefschriftMaken | proefschriftmaken.nl
Cover design:	Wendy Schoneveld
ISBN:	978-94-6534-383-9

The printing of this thesis was financially supported by Scholten-Cordes Fonds and Spaarne Gasthuis.

Optimizing Care in Lumbar Radiculopathy and Neurogenic Claudication

From injection to inference, and from clinician to algorithm

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 30 juni 2026
klokke 14:30 uur

door

Eduard Jacques Alfons Verheijen

geboren te 's-Gravenhage
in 1993

Promotoren

Prof. dr. C.L.A. Vleggeert-Lankamp
Prof. dr. ir. M. Staring

Leden promotiecommissie

Prof. dr. J. Gussekloo
Prof. dr. ir. B.F.W. Lelieveldt
Prof. dr. R.H.M.A. Bartels, Radboud Universiteit Nijmegen
Dr. H. Arnts, Radboud Universiteit Nijmegen

TABLE OF CONTENTS

Part I

Optimizing care for lumbar radiculopathy

1

General introduction and outline

Lumbosacral radiculopathy, often used interchangeably with sciatica in the literature despite referring more broadly to a condition of the lumbar or sacral nerve roots, is a neurological disorder with a reported lifetime prevalence of up to 43% [1]. Generally, it is characterized by unilateral radicular pain extending along one of the exiting nerve roots and is caused by either lumbosacral disc herniation (LDH) or degenerative spinal stenosis (LSS) due to age-related disc degeneration, hypertrophy of the ligamentum flavum and zygapophyseal joint, and osteophyte formation [2]. Radicular pain is frequently accompanied by functional impairments, back pain and motor-sensory deficits [3]. Originally, nerve root compression was assumed to be the primary cause of symptoms. However, research over the past decades has demonstrated that a significant proportion of patients exhibit severe symptoms despite the absence of nerve root compression on imaging. Conversely, some patients present with clear radiological evidence of nerve root impingement but experience little to no symptoms [4-6]. As a result, the general idea has changed and it is now accepted that the pathophysiology of lumbosacral radiculopathy is a multifaceted interplay involving inflammatory processes and immunological responses in conjunction with mechanical compression [7-10]. Exposure of the nerve root to nucleus pulposus material has been shown to elicit an inflammatory response, including cytokine release and recruitment of pro-inflammatory cells, which exacerbates pain and neural dysfunction. Stenosis can result in nerve root oedema and inflammation caused by local ischemia as a result from congestion of venous blood around the nerve root inducing cytokine release. Additionally, vertebral end plate devascularization may strengthen this response [10-13].

Disease progression, however, usually differs as patients with LDH are younger on average and have a more acute onset of symptoms which may aggravate when coughing, sneezing or leg-straightening, while patients with LSS often are older and experience gradually increasing symptoms that may exacerbate during standing or walking. However, for both patient groups, the impact on daily functioning can be severe, necessitating effective management strategies [14]. Since lumbosacral radiculopathy due to LDH may follow a favorable, self-limiting course, and patient outcomes after one year of follow-up have been shown to be similar between surgical and non-surgical groups, initial treatment is conservative for several weeks to months before considering imaging and non-conservative interventions [15-18]. There is less consensus on the optimal treatment approach for LSS, as its symptoms usually follow a more chronic course, though some patients may still experience spontaneous resolution of their symptoms [19-21]. In both patient groups, oral pain medication often provides inadequate relief, and the substantial impact of the condition leads to significant medical and socio-economic costs [22].

One of the primary interventional treatments for lumbosacral radiculopathy is an epidural steroid injection (ESI), which aims to mitigate inflammation and provide symptomatic relief. The origins of this technique date back to 1895 when Bier anesthetized the lower body of one of his residents by injecting a cocaine solution into the intrathecal space [23]. Building on this, Cathelin and Sicard in the early 1900s adapted the approach by using the caudal route to remain within the epidural space [24,25], and it was not until the 1950s that Robecchi and Capra introduced the first periradicular steroid injection with hydrocortisone, laying the foundation for modern ESI [26]. In addition to the caudal approach, interlaminar and transforaminal techniques have been developed. Particularly the transforaminal epidural steroid injection (TEI) approach is now widely employed as it is considered the most selective one of techniques to deliver corticosteroids near the affected nerve root.

Given its potential to reduce inflammatory mediators and modulate immune responses [11,27-29], TEI is commonly administered as a minimally invasive alternative for patients with lumbar radiculopathy who do not respond adequately to conservative treatment [30]. However, its efficacy remains a topic of debate due to inconsistent findings in the literature. While some studies report significant pain relief and functional improvement, others suggest only limited benefits compared to placebo treatments. A 2012 systematic review and meta-analysis by Pinto et al. found a slight favour for corticosteroids but questioned the clinical utility [31]. Since then, several additional randomized controlled trials have been conducted, specifically for the transforaminal approach, highlighting the need for a renewed evaluation of the effectiveness of TEI.

The effectiveness of TEI in patients with lumbosacral radiculopathy may depend on the underlying etiology; however, this remains uncertain. The majority of studies have focused on patients with MRI-confirmed LDH, while relatively few have evaluated TEI outcomes in those with radiculopathy due to other etiologies or in the absence of a compressive lesion on imaging [32,33]. The studies that have compared the effect of TEI between patients with LDH and those with other etiologies suggest comparable outcomes, though findings vary based on the outcome measure and duration of follow-up [34-36]. If TEI provides symptom relief independent of the underlying pathology, routine pre-treatment MRI scans may not be necessary. One study specifically indicated that MRI prior to TEI may be redundant, as it did not significantly influence treatment effectiveness or clinical decision-making [36]. However, the overall body of evidence remains constrained by the small number of studies and the lack of prospective data.

Beyond the underlying etiology of symptoms, no consensus exists regarding other factors associated with TEI effectiveness. Some studies suggest that patients with a shorter duration of symptoms may experience greater benefit from TEI compared to those with a more chronic history [32], or that early pain relief following TEI could serve as a prognostic indicator of long-term outcomes [37,38]. Various demographic, clinical, and radiological variables have been explored; however, no strong conclusions have been drawn, largely due to the considerable heterogeneity in patient populations, epidural steroid techniques, and outcome measures. This lack of consistency has precluded the development of a predictive model for TEI treatment success and, consequently, the implementation of more patient-tailored therapeutic strategies for patients suffering from lumbosacral radiculopathy.

AIMS AND OUTLINE OF THIS THESIS:

1. The only systematic review and meta-analysis comparing the effectiveness of a transforaminal epidural steroid injection (TEI) to placebo – either a transforaminal injection with saline or a sham-injection in the para-spinal muscle – dates back to 2012 [31]. Since then, multiple randomized controlled trials have been conducted, primarily evaluating transforaminal injections. Given these new studies, a reassessment is warranted to determine whether the additional data alter the existing conclusions. Therefore, this first aim is to systematically review and update the current evidence on TEI efficacy compared to placebo in patients with lumbosacral radiculopathy (Chapter 2).
2. TEI has become an increasingly popular treatment for patients with lumbosacral radiculopathy. Since this condition can be caused by a variety of spinal pathology, it is essential to determine whether TEI effectiveness depends on the underlying etiology. Hence, the second aim is to assess the effectiveness of TEI in patients with MRI-confirmed lumbar disc herniation (LDH) compared to those with alternative causes or no compressive lesion on imaging, using data from a large retrospective cohort (Chapter 3).
3. To optimize treatment strategies, it is essential to identify factors that positively or negatively affect patient outcome after treatment with TEI. Therefore, the third aim is to comprehensively review all studies that have examined demographic, clinical, or radiological parameters and their association with the outcome of a single transforaminal injection in patients with lumbar radiculopathy secondary to disc herniation or spinal stenosis (Chapter 4).
4. Current literature suggests that patients with a shorter duration of symptoms may benefit more from TEI. Moreover, there is some evidence that

routine MRI before TEI may be redundant. At present, patients are treated conservatively for several weeks to months before referral to a neurologist and subsequent MRI examination. It is usually only after these steps that they become eligible for TEI or surgical intervention. During this period, symptoms often remain inadequately managed, impairing daily function and contributing to absenteeism. Given these findings, our fourth aim is to investigate whether patients may benefit from early TEI administration. To address this, we present the protocol for a randomized controlled trial comparing the effectiveness of TEI to usual care in patients experiencing acute lumbar radiculopathy symptoms (Chapter 5).

5. The identification of prognostic factors for TEI is significantly limited by the lack of large cohorts and prospective data. Furthermore, substantial heterogeneity in study designs and treatment protocols—combining outcomes from caudal, interlaminar, and transforaminal injections using various corticosteroids—complicates personalizing treatment strategies. To achieve this, the fifth aim is to develop a prediction model for TEI success. We present the protocol for a large, multi-center, prospective cohort study designed to assess the effects of TEI in patients with MRI-confirmed disc herniation or stenosis, congruent with clinical findings, and to identify demographic, clinical and radiological parameters associated with treatment outcomes (Chapter 6).
6. Vertebral endplate changes, known as Modic changes (MC), have been proposed as markers of inflammatory processes and are associated with disc herniation and less favorable outcomes after surgery [39-46]. It has been suggested that the absence of MC correlates with reduced radicular pain and improved clinical outcomes, whereas their presence may aggravate symptoms [46]. Consequently, in patients with MC at the level of disc herniation, inflammation may play a more significant role, and TEI may be more effective. The sixth aim is to assess this association using prospective data from our large cohort study (Chapter 7).

REFERENCES

1. Konstantinou K, Dunn KM (2008) Sciatica: review of epidemiological studies and prevalence estimates. Spine (Phila Pa 1976) 33 (22):2464-2472. doi:10.1097/BRS.0b013e318183a4a2
2. Waldman SD (2019) 84 - Spinal Stenosis. In: Waldman SD (ed) Atlas of Common Pain Syndromes (Fourth Edition). Elsevier, Philadelphia, pp 324-327. doi:https://doi.org/10.1016/B978-0-323-54731-4.00084-0
3. Koes BW, van Tulder MW, Peul WC (2007) Diagnosis and treatment of sciatica. Bmj 334 (7607):1313-1317. doi:10.1136/bmj.39223.428495.BE
4. el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, Van der Kallen BF, van den Hout WB, Jacobs WC, Koes BW, Peul WC (2013) Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med 368 (11):999-1007. doi:10.1056/NEJMoa1209250
5. van Rijn JC, Klemetso N, Reitsma JB, Majoie CB, Hulsmans FJ, Peul WC, Bossuyt PM, Heeten GJ, Stam J (2006) Symptomatic and asymptomatic abnormalities in patients with lumbosacral radicular syndrome: Clinical examination compared with MRI. Clin Neurol Neurosurg 108 (6):553-557. doi:10.1016/j.clineuro.2005.10.003
6. Karppinen J, Malmivaara A, Tervonen O, Pääkkö E, Kurunlahti M, Syrjälä P, Vasari P, Vanharanta H (2001) Severity of symptoms and signs in relation to magnetic resonance imaging findings among sciatic patients. Spine (Phila Pa 1976) 26 (7):E149-154. doi:10.1097/00007632-200104010-00015
7. Stafford MA, Peng P, Hill DA (2007) Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth 99 (4):461-473. doi:10.1093/bja/aem238
8. Wang K, Bao JP, Yang S, Hong X, Liu L, Xie XH, Wu XT (2016) A cohort study comparing the serum levels of pro- or anti-inflammatory cytokines in patients with lumbar radicular pain and healthy subjects. Eur Spine J 25 (5):1428-1434. doi:10.1007/s00586-015-4349-4
9. Valat JP, Genevay S, Marty M, Rozenberg S, Koes B (2010) Sciatica. Best Pract Res Clin Rheumatol 24 (2):241-252. doi:10.1016/j.berh.2009.11.005
10. Dower A, Davies MA, Ghahreman A (2019) Pathologic Basis of Lumbar Radicular Pain. World Neurosurg 128:114-121. doi:10.1016/j.wneu.2019.04.147
11. McLain RF, Kapural L, Mekhail NA (2005) Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J 5 (2):191-201. doi:10.1016/j.spinee.2004.10.046
12. Mulleman D, Mammou S, Griffoul I, Watier H, Goupille P (2006) Pathophysiology of disk-related sciatica. I.--Evidence supporting a chemical component. Joint Bone Spine 73 (2):151-158. doi:10.1016/j.jbspin.2005.03.003
13. Lee BH, Moon SH, Suk KS, Kim HS, Yang JH, Lee HM (2020) Lumbar Spinal Stenosis: Pathophysiology and Treatment Principle: A Narrative Review. Asian Spine J 14 (5):682-693. doi:10.31616/asj.2020.0472
14. Deyo RA, Mirza SK (2016) CLINICAL PRACTICE. Herniated Lumbar Intervertebral Disk. N Engl J Med 374 (18):1763-1772. doi:10.1056/NEJMcp1512658
15. Chiu CC, Chuang TY, Chang KH, Wu CH, Lin PW, Hsu WY (2015) The probability of spontaneous regression of lumbar herniated disc: a systematic review. Clin Rehabil 29 (2):184-195. doi:10.1177/0269215514540919

16. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, Thomeer RT, Koes BW (2007) Surgery versus prolonged conservative treatment for sciatica. N Engl J Med 356 (22):2245-2256. doi:10.1056/NEJMoa064039
17. Kreiner DS, Hwang SW, Easa JE, Resnick DK, Baisden JL, Bess S, Cho CH, DePalma MJ, Dougherty P, 2nd, Fernand R, Ghiselli G, Hanna AS, Lamer T, Lisi AJ, Mazanec DJ, Meagher RJ, Nucci RC, Patel RD, Sembrano JN, Sharma AK, Summers JT, Taleghani CK, Tontz WL, Jr., Toton JF (2014) An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. Spine J 14 (1):180-191. doi:10.1016/j.spinee.2013.08.003
18. Lorio M, Kim C, Araghi A, Inzana J, Yue JJ (2020) International Society for the Advancement of Spine Surgery Policy 2019-Surgical Treatment of Lumbar Disc Herniation with Radiculopathy. Int J Spine Surg 14 (1):1-17. doi:10.14444/7001
19. Anderson DB, Luca K, Jensen RK, Eyles JP, Van Gelder JM, Friedly JL, Maher CG, Ferreira ML (2021) A critical appraisal of clinical practice guidelines for the treatment of lumbar spinal stenosis. Spine J 21 (3):455-464. doi:10.1016/j.spinee.2020.10.022
20. Katz JN, Zimmerman ZE, Mass H, Makhni MC (2022) Diagnosis and Management of Lumbar Spinal Stenosis: A Review. Jama 327 (17):1688-1699. doi:10.1001/jama.2022.5921
21. Kreiner DS, Shaffer WO, Baisden JL, Gilbert TJ, Summers JT, Toton JF, Hwang SW, Mendel RC, Reitman CA (2013) An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). Spine J 13 (7):734-743. doi:10.1016/j.spinee.2012.11.059
22. Lambeek LC, van Tulder MW, Swinkels IC, Koppes LL, Anema JR, van Mechelen W (2011) The trend in total cost of back pain in The Netherlands in the period 2002 to 2007. Spine (Phila Pa 1976) 36 (13):1050-1058. doi:10.1097/BRS.0b013e3181e70488
23. Bier A (1899) Versuche über Cocainisirung des Rückenmarkes. Deutsche Zeitschrift für Chirurgie 51 (3):361-369. doi:10.1007/BF02792160
24. Sicard A (1901) Les injections medicamenteuses extra-durales par voie sacrococcygienne. Compt Rend Soc De Biol 53:396-398
25. Cathelin F (1901) Une nouvelle voie d'injection rachidienne: methode des injections epidurales par le procede du canal sacre-applications a l'homme. Compt Rend Soc De Biol 53:452-453
26. Robecchi A, Capra R (1952) [Hydrocortisone (compound F); first clinical experiments in the field of rheumatology]. Minerva Med 43 (98):1259-1263
27. Kantrowitz F, Robinson DR, McGuire MB, Levine L (1975) Corticosteroids inhibit prostaglandin production by rheumatiod synovia. Nature 258 (5537):737-739. doi:10.1038/258737a0
28. Johansson A, Hao J, Sjölund B (1990) Local corticosteroid application blocks transmission in normal nociceptive C-fibres. Acta Anaesthesiol Scand 34 (5):335-338. doi:10.1111/j.1399-6576.1990.tb03097.x
29. Colligan N, Gupta S (2010) Epidural steroids. Continuing Education in Anaesthesia, Critical Care and Pain 10 (1):1-5. doi:10.1093/bjaceaccp/mkp043
30. Manchikanti L, Falco FJ, Singh V, Pampati V, Parr AT, Benyamin RM, Fellows B, Hirsch JA (2012) Utilization of interventional techniques in managing chronic pain in the Medicare population: analysis of growth patterns from 2000 to 2011. Pain Physician 15 (6):E969-982

31. Pinto RZ, Maher CG, Ferreira ML, Hancock M, Oliveira VC, McLachlan AJ, Koes B, Ferreira PH (2012) Epidural corticosteroid injections in the management of sciatica: a systematic review and meta-analysis. Ann Intern Med 157 (12):865-877. doi:10.7326/0003-4819-157-12-201212180-00564
32. MacVicar J, King W, Landers MH, Bogduk N (2013) The effectiveness of lumbar transforaminal injection of steroids: a comprehensive review with systematic analysis of the published data. Pain Med 14 (1):14-28. doi:10.1111/j.1526-4637.2012.01508.x
33. Smith CC, McCormick ZL, Mattie R, MacVicar J, Duszynski B, Stojanovic MP (2020) The Effectiveness of Lumbar Transforaminal Injection of Steroid for the Treatment of Radicular Pain: A Comprehensive Review of the Published Data. Pain Med 21 (3):472-487. doi:10.1093/pm/pnz160
34. Maus TP, El-Yahchouchi CA, Geske JR, Carter RE, Kaufmann TJ, Wald JT, Diehn FE (2016) Imaging Determinants of Clinical Effectiveness of Lumbar Transforaminal Epidural Steroid Injections. Pain Med 17 (12):2176-2184. doi:10.1093/pm/pnw043
35. Ng LC, Sell P (2004) Outcomes of a prospective cohort study on peri-radicular infiltration for radicular pain in patients with lumbar disc herniation and spinal stenosis. Eur Spine J 13 (4):325-329. doi:10.1007/s00586-003-0649-1
36. Cohen SP, Gupta A, Strassels SA, Christo PJ, Erdek MA, Griffith SR, Kurihara C, Buckenmaier CC, 3rd, Cornblath D, Vu TN (2012) Effect of MRI on treatment results or decision making in patients with lumbosacral radiculopathy referred for epidural steroid injections: a multicenter, randomized controlled trial. Arch Intern Med 172 (2):134-142. doi:10.1001/archinternmed.2011.593
37. El-Yahchouchi C, Wald J, Brault J, Geske J, Hagen C, Murthy N, Kaufmann T, Thielen K, Morris J, Diehn F, Amrami K, Carter R, Shelerud R, Maus T (2014) Lumbar transforaminal epidural steroid injections: does immediate post-procedure pain response predict longer term effectiveness? Pain Med 15 (6):921-928. doi:10.1111/pme.12347
38. Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN (2016) The Value of Short-Term Pain Relief in Predicting the One-Month Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg 96:323-333. doi:10.1016/j.wneu.2016.09.016
39. Dudli S, Miller S, Demir-Deviren S, Lotz JC (2018) Inflammatory response of disc cells against Propionibacterium acnes depends on the presence of lumbar Modic changes. Eur Spine J 27 (5):1013-1020. doi:10.1007/s00586-017-5291-4
40. Viswanathan VK, Shetty AP, Rajasekaran S (2020) Modic changes - An evidence-based, narrative review on its patho-physiology, clinical significance and role in chronic low back pain. J Clin Orthop Trauma 11 (5):761-769. doi:10.1016/j.jcot.2020.06.025
41. Heggli I, Laux CJ, Mengis T, Karol A, Cornaz F, Herger N, Aradi-Vegh B, Widmer J, Burkhard MD, Farshad-Amacker NA, Pfammatter S, Wolski WE, Brunner F, Distler O, Farshad M, Dudli S (2023) Modic type 2 changes are fibroinflammatory changes with complement system involvement adjacent to degenerated vertebral endplates. JOR Spine 6 (1):e1237. doi:10.1002/jsp2.1237
42. Dudli S, Sing DC, Hu SS, Berven SH, Burch S, Deviren V, Cheng I, Tay BKB, Alamin TF, Ith MAM, Pietras EM, Lotz JC (2017) ISSLS PRIZE IN BASIC SCIENCE 2017: Intervertebral disc/bone marrow cross-talk with Modic changes. Eur Spine J 26 (5):1362-1373. doi:10.1007/s00586-017-4955-4

43. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR (1988) Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology 166 (1 Pt 1):193-199. doi:10.1148/radiology.166.1.3336678
44. Dudli S, Karol A, Giudici L, Heggli I, Laux CJ, Spirig JM, Wanivenhaus F, Betz M, Germann C, Farshad-Amacker N, Brunner F, Distler O, Farshad M (2022) CD90-positive stromal cells associate with inflammatory and fibrotic changes in modic changes. Osteoarthr Cartil Open 4 (3):100287. doi:10.1016/j.ocarto.2022.100287
45. Albert HB, Kjaer P, Jensen TS, Sorensen JS, Bendix T, Manniche C (2008) Modic changes, possible causes and relation to low back pain. Med Hypotheses 70 (2):361-368. doi:10.1016/j.mehy.2007.05.014
46. Djuric N, Yang X, Ostelo R, van Duinen SG, Lycklama À Nijeholt GJ, van der Kallen BFW, Peul WC, Vleggeert-Lankamp CLA (2019) Disc inflammation and Modic changes show an interaction effect on recovery after surgery for lumbar disc herniation. Eur Spine J 28 (11):2579-2587. doi:10.1007/s00586-019-06108-9

2

Epidural steroid compared to placebo injection in sciatica: a systematic review and meta-analysis

E.J.A. Verheijen, C.A. Bonke, E.M.J. Amorij, C.L.A. Vleggeert-Lankamp

Eur Spine J 2021

ABSTRACT

Purpose

The purpose of this systematic review and meta-analysis was to determine whether epidural steroid injections (ESI) are superior to epidural or non-epidural placebo injections in sciatica patients.

Methods

The PubMed, Embase, Cochrane Library, and Web of science databases were searched for trials comparing ESI to epidural or non-epidural placebo. Risk of bias was assessed using the Cochrane RoB 2 tool. The primary outcome measures were pooled using a random-effects model for 6-week, 3-month, and 6-month follow-up. Secondary outcomes were described qualitatively. Quality of evidence was graded using GRADE classification.

Results

Seventeen out of 732 articles were included. ESI was superior compared to epidural placebo at 6 weeks (− 8.6 [− 13.4; − 3.9]) and 3 months (− 5.2 [− 10.1; − 0.2]) for leg pain and at 6 weeks for functional status (− 4.1 [− 6.5; − 1.6]), though the minimally clinical important difference (MCID) was not met. There was no difference in ESI and placebo for back pain, except for non-epidural placebo at 3 months (6.9 [1.3; 12.5]). Proportions of treatment success were not different. ESI reduced analgesic intake in some studies and complication rates are low.

Conclusion

The literature indicates that ESI induces larger improvements in pain and disability on the short term compared to epidural placebo, though evidence is of low to moderate quality and MCID is not met. Strong conclusions for longer follow-up or for comparisons with non-epidural placebo cannot be drawn due to general low quality of evidence and limited number of studies. Epidural injections can be considered a safe therapy.

INTRODUCTION

Sciatica is a common spinal condition with high reported lifetime prevalence and is generally caused by a lumbosacral disc herniation (LDH) [1]. Patients usually present with unilateral leg pain with ensuing disabilities. Other associated clinical manifestations include back pain, motor-sensory deficits, and reflex abnormalities [2]. Despite the debilitating physical burden, sciatica has a favourable prognosis due to its self-limiting essence and hence, most patients are initially treated conservatively [3].

It is assumed that sciatica symptoms are triggered by a complex interaction of compression-related, inflammatory, and immunological mechanisms [4]. Physical impingement of a nerve root from LDH is not necessarily sufficient to induce pain, as a substantial group of patients presents with neural compromise on imaging in absence of clinical symptoms and vice versa [5–7]. Possibly, in addition to nerve root compression, immunological and inflammatory processes play a key role. Exposure to nucleus pulposus tissue is assumed to cause an auto-immune response leading to cytokine production and involvement of pro-inflammatory cells. Additionally, vertebral end plate devascularization may strengthen this response [8–12].

These inflammatory processes are the primary target of epidural steroid injection (ESI) treatment. Through an interlaminar (IL), transforaminal (TF) or caudal approach anti-inflammatory medication can be deposited in close proximity of the affected nerve root, which is presumed to inhibit production of inflammatory mediators and to downregulate the immunological response. Subsequently, inflammation is decreased resulting in pain reduction and functional improvement for the patient [8, 13–15]. Although several studies have investigated the efficacy of ESI in comparison with placebo, they have generated inconsistent results precluding an unequivocal recommendation on ESI therapy. However, despite the lack of consensus on efficacy, this treatment has been firmly established as a minimally invasive method for pain management in sciatica with continuously increasing utilization rates [16–18]. Therefore, this review explores the validity of ESI treatment compared to epidural and non-epidural placebo in sciatica patients in current practice.

METHODS

This systematic review and meta-analysis were conducted in accordance with the PRISMA guidelines.

Search and selection

The PubMed, Embase, Cochrane Library and Web of Science databases were searched on August 20, 2020, using an all-encompassing search strategy constructed by an expert librarian. The search strategy combined strings for randomized-controlled trials with sciatica patients, treatment with ESI compared to epidural or non-epidural placebo and appropriate outcome measures (ESM 1). Retrieved studies were selected first on title and abstract by three independent reviewers (EV, CB, EA). Consequently, selected studies and previously published systematic reviews were subjected to citation tracking and all obtained articles were reviewed in full text. In case of a discrepancy, consensus was reached through discussion or consultation of a fourth reviewer (CVL).

Inclusion and exclusion criteria

Articles were eligible if they described an RCT that compared injection of steroid into the epidural space with injection of placebo using the same technique or with non-epidural placebo. All three techniques (caudal, IL and TF) were accepted. Epidural placebo was defined either as an inert substance without pharmacological activity (e.g. saline) or as a short-living local anaesthetic (e.g. lidocaine) delivered to the epidural space. Non-epidural placebo was defined as an inert substance without pharmacological activity administered into soft tissue surrounding the lumbar spine (e.g. subcutaneous). Studies were eligible if they provided data on sciatica patients, unless they only reported specifically on patients with a stenosis, or if they provided data separately for a subgroup of sciatica patients without stenosis. Studies were included if treatment efficacy was assessed using a validated instrument for pain or disability in at least 20 patients for a minimum follow-up of 2 weeks. Studies that evaluated pain without specifying whether this was leg pain were eligible as well, since leg pain is usually worse than back pain in sciatica patients. For assessment of pain, the visual analogue scale (VAS) and numerical rating scale (NRS) and for disability, the Oswestry Disability Index (ODI) and Roland–Morris Disability Questionnaire (RMDQ) were considered appropriate instruments. Only reports in English, Dutch, German, French, or Spanish were accepted.

Risk-of-bias assessment

The three reviewers individually assessed the risk of bias of included articles using the Cochrane risk of bias tool (RoB 2) [19]. For the second domain, the effect of assignment to intervention was determined. In addition to the published article, trial registrations and protocols were used if available to assess risk-of-bias questions, or the corresponding author was contacted for clarification. Differences between answers to the questions were resolved during a consensus meeting. Data extraction and analysis Descriptive and quantitative data were retrieved from each study by two reviewers independently with the third reviewer verifying the final data extraction sheet. Information was collected regarding authors, publication date, patients, interventions, primary and secondary outcomes and results. Leg pain, back pain and disability were assessed as the primary patient outcomes. Pain medication use and complications were assessed as secondary outcomes. For continuous outcome data, means and standard deviations (SD) and mean differences (MD) with corresponding standard error (SE) and 95% confidence interval (CI) adjusted for baseline differences were collected. Alternatively, unadjusted MDs and SE were calculated preferring the use of change scores to final values [20]. If SD was missing, the value was imputed preferably using SD values from the same study or else from a comparable study. For dichotomous outcome data, absolute numbers, percentages, risk ratio (RR) or odds ratio (OR) with SE or 95% CI were obtained. In case outcome data were only presented graphically, numeric data was extracted from the figure. Continuous data were converted to a 0–100 scale for comparability purposes. Treatment arms of the same type within a study were combined (e.g. three steroid groups were merged) and analysed as a single intervention group [20].

Meta-analysis using random-effects model in R (R Foundation for Statistical Computing, Vienna, Austria) was performed for pooling of primary outcomes if patient groups were considered sufficiently clinically homogeneous. Results were pooled separately for each epidural technique and together combining all three approaches for three follow-up timeframes: 6 weeks, 3 months and 6 months. Sensitivity analyses were performed for the combined pooled estimate. Sensitivity analysis for heterogeneity (I2) was conducted when I2 > 25%. Publication bias was assessed using funnel plots and Egger's test but only discussed if assessment included at least five studies [21]. Quality of evidence for pooled results was graded using the GRADE system [22]. Secondary outcomes were assessed qualitatively. Detailed study data and figures are available as Electronic Supplementary Material (ESM).

RESULTS

Article selection

The search yielded a total of 413 unique references of which 57 articles were selected for full-text assessment. Ultimately, 17 reports [15, 23–38] were included for this review (Fig. 1). Thirteen articles focussed on epidural placebo [15, 24–27, 29, 32–38], two on non-epidural placebo [23, 30] and two included both placebo groups [28, 31]. A more elaborate overview of study characteristics is given in ESM 2.

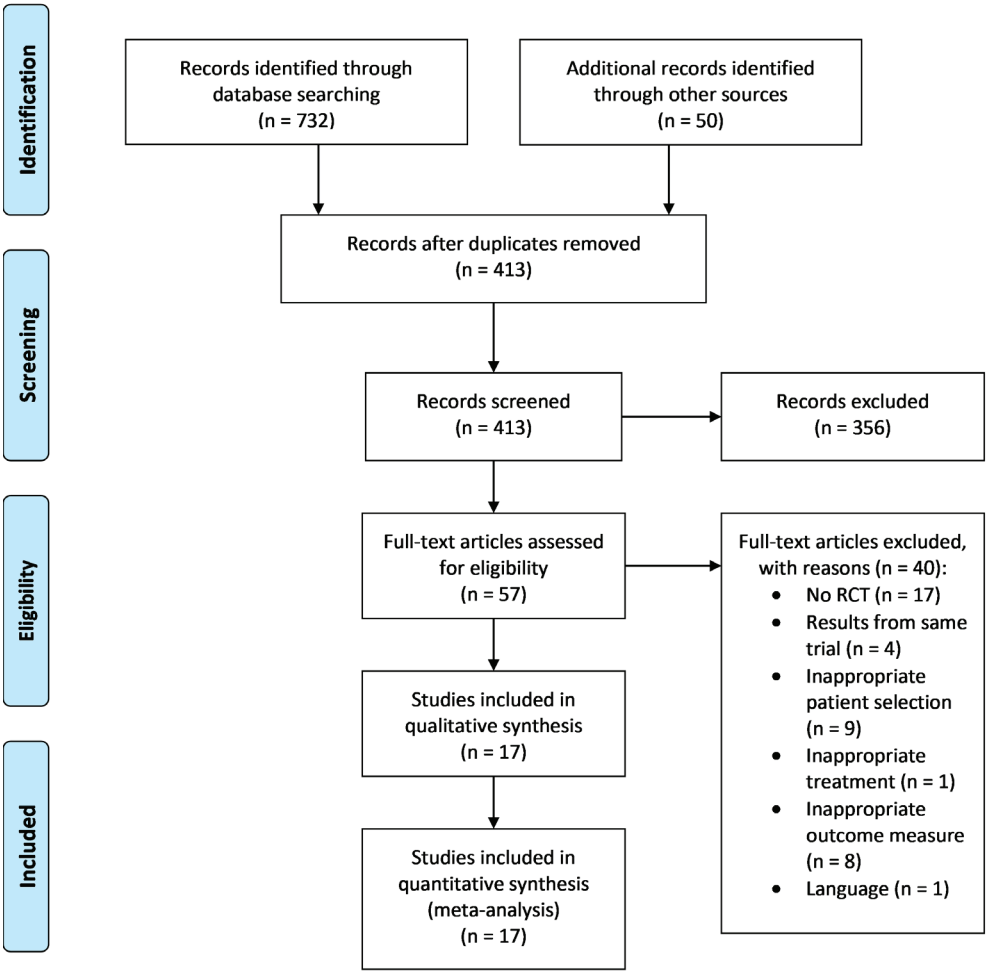


Fig. 1 Flowchart of the article search and selection process

Risk-of-bias assessment

Of the 17 trials, five were considered low risk of bias [23, 26, 28, 31, 32], two raised some concerns [15, 30] and ten studies were scored as high-risk [24, 25, 27, 29, 33–38]. Initially, the five low-risk reports were categorized as raising some concerns only due to unavailability of a pre-specified statistical analysis plan. Since a statistical analysis plan was unavailable for all studies, these five trials were judged as low risk (ESM 3).

An overview of study data used for meta-analysis is given in ESM 4. Publication bias data is presented in ESM 10.

Treatment effect on leg pain

Epidural steroid versus epidural placebo

For 6-week and for 3-month follow-up, the pooled estimate favoured ESI to epidural placebo for reduction of leg pain. Pooling for 6-month follow-up only demonstrated a statistically significant difference after exclusion of three studies [29, 32, 36] for heterogeneity (Table 1) (ESM 7). Assessment of publication bias was inconclusive for 3- and 6-month follow-up and showed an absence for 6-week follow-up. Sensitivity analyses demonstrated various changes in direction of the pooled estimate for all follow-up timeframes with the most pronounced shift being between sciatica patients with a radiological and patients with a clinical diagnosis for 6-week follow-up (ESM 11).

GRADE quality of evidence: low (ESM 12).

Table 1 Pooled effect estimates from continuous data for leg pain and adjusted for sensitivity analysis of heterogeneity

	Primary analysis					***Sensitivity analysis for heterogeneity***				
	Number of studies	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**	**Number of studies**	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**
Epidural steroid vs. epidural placebo										
6-week FU	10	997	-8.6 (-13.4; -3.9)	<0.01	70%	7		-5.9 (-8.7; -3.2)	<0.01	14%
3-month FU	10	1188	-5.2 (-10.1; -0.2)	0.04	83%	7		-6.8 (-10.3; -3.2)	<0.01	25%
6-month FU	7	677	-2.7 (-8.0; 2.6)	0.31	75%	4		-5.1 (-8.0; -2.3)	<0.01	0%

Table 1 Pooled effect estimates from continuous data for leg pain and adjusted for sensitivity analysis of heterogeneity (*continued*)

	Primary analysis					***Sensitivity analysis for heterogeneity***				
	Number of studies	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**	**Number of studies**	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**
Epidural steroid vs. non-epidural placebo										
6-week FU	4	399	-8.1 (-17.8; 1.6)	0.10	80%	2		-0.1 (-3.9; 3.7)	0.97	0%
3-month FU	3	337	-1.0 (-17.9; 15.8)	0.90	92%	*	-	-	-	-
6-month FU	2	294	1.7 (-2.1; 5.4)	0.38	0%	*	-	-	-	-

A negative MD indicates a favour for ESI; a positive MD indicates a favour for placebo

CI confidence interval, FU follow-up, MD mean difference

* The limited number of studies did not allow for sensitivity analysis of heterogeneity

Epidural steroid versus non-epidural placebo

When comparing ESI to non-epidural placebo, the pooled estimate was non-significant for all three follow-up periods (Table 1) (ESM 7). Sensitivity analyses could only be performed for 6-week and 3-month follow-up and demonstrated large differences between effect estimates (ESM 11).

GRADE quality of evidence: low (6-week and 6-month) and moderate (3-month) (ESM 12).

Treatment effect on back pain

Epidural steroid versus epidural placebo

The pooled effect estimate was not significantly different between ESI and epidural placebo for all follow-up time frames (Table 2) (ESM 8). Sensitivity analyses affected the 6-week pooled result only slightly, whereas for 3-month follow-up larger variations were observed. For 6-month follow-up, sensitivity analysis showed favour for ESI when fluoroscopic image guidance was used and in patients with (sub) acute symptoms (ESM 11).

GRADE quality of evidence: moderate (6-week), very low (3-month) and low (6-month) (ESM 12).

Table 2 Pooled effect estimates from continuous data for back pain

	Primary analysis				
	Number of studies	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**
Epidural steroid vs. epidural placebo					
6-week FU	3	290	-2.9 (-6.8; 0.9)	0.14	0%
3-month FU	2	227	0.7 (-23.5; 25.0)	0.95	94%
6-month FU	2	225	-4.9 (19.9; 10.2)	0.53	84%
Epidural steroid vs. non-epidural placebo					
6-week FU	2	302	-1.7 (-6.6; 3.1)	0.49	34%
3-month FU	2	298	6.9 (1.3; 12.5)	0.02	42%
6-month FU	2	294	1.3 (-2.2; 4.9)	0.46	0%

A negative MD indicates a favour for ESI; a positive MD indicates a favour for placebo. Sensitivity analysis of heterogeneity was not feasible for all follow-up time-frames due to the limited number of studies.

CI confidence interval, FU follow-up, MD mean difference

* The limited number of studies did not allow for sensitivity analysis of heterogeneity

Epidural steroid versus non-epidural placebo

The pooled effect estimate was only statistically significant for 3-month follow-up in favour of non-epidural placebo (Table 2) (ESM 8). Sensitivity analyses were not feasible due to the limited number of studies included.

GRADE quality of evidence: moderate (ESM 12).

Treatment effect on functional status

Epidural steroid versus epidural placebo

ESI was favoured to placebo for improvement of disability at 6-week follow-up, and at 3-month follow-up after

adjustment for heterogenic studies [31]. For 6-month follow-up, none of the interventions was favoured (Table 3) (ESM 9). Assessment of publication bias was inconclusive. Sensitivity analyses resulted in different pooled estimates at 3 and 6-month follow-up (ESM 11).

GRADE quality of evidence: moderate (6-week), low (3-month) and very low (6-month) (ESM 12).

Table 3 Pooled effect estimates from continuous data for functional status and adjusted for sensitivity analysis of heterogeneity

	Primary analysis					*Sensitivity analysis for heterogeneity*				
	Number of studies	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**	**Number of studies**	**Number of patients**	**MD (95% CI)**	**P-value**	**I^2**
Epidural steroid vs. epidural placebo										
6-week FU	6	624	-4.1 (-6.5; -1.6)	<0.01	35%	5		-2.5 (-4.5; -0.5)	0.01	0%
3-month FU	9	981	-2.5 (-5.5; 0.5)	0.10	71%	8		-4.1 (-5.9; -2.3)	<0.01	0%
6-month FU	6	653	-1.0 (-5.4; 3.5)	0.67	84%	2		-2.6 (-6.1; 0.8)	0.14	0%
Epidural steroid vs. non-epidural placebo										
6-week FU	2	302	-0.8 (-3.3; 1.6)	0.52	25%	*	-	-	-	-
3-month FU	2	298	4.0 (-3.0; 11.0)	0.26	83%	*	-	-	-	-
6-month FU	2	294	2.8 (-3.9; 9.5)	0.41	84%	*	-	-	-	-

A negative MD indicates a favour for ESI; a positive MD indicates a favour for placebo.

CI: confidence interval; FU: follow-up; MD: mean difference

* The limited number of studies did not allow for sensitivity analysis of heterogeneity

Epidural steroid versus non-epidural placebo

The effect estimate favoured none of the interventions for 6-week, 3- and 6-month follow-up (Table 3) (ESM 9). Sensitivity analyses were not feasible due to the limited number of studies included.

GRADE quality of evidence: moderate (ESM 12).

Proportions of treatment success

For studies with treatment success defined as ≥ 50% improvement in pain scores, neither ESI nor epidural pla-

cebo were favoured at 6-week, 3-month and 6-month follow-up (Table 4) (ESM 7). Sensitivity analyses produced varying results with mostly minor differences in effect estimates (ESM 11).

GRADE quality of evidence: high (6-week) and very low (3-month and 6-month) (ESM 12).

Table 4 Pooled effect estimates from proportional data for leg pain and functional status, and adjusted for sensitivity analysis of heterogeneity

	Primary analysis					*Sensitivity analysis for heterogeneity*				
	Number of studies	**Number of patients**	**RR (95% CI)**	**P-value**	**I^2**	**Number of studies**	**Number of patients**	**RR (95% CI)**	**P-value**	**I^2**
≥50% improvement in pain scores										
6-week FU	2	150	2.3 (0.9; 5.9)	0.08	81%	*	-	-	-	-
3-month FU	5	487	1.1 (1.0; 1.3)	0.15	51%	4	418	1.1 (1.0; 1.2)	0.29	0%
6-month FU	5	487	1.1 (0.9; -1.3)	0.24	60%	2	178	1.0 (0.7; 1.4)	0.98	24%
≥50% improvement in ODI scores										
3-month FU	3	360	1.1 (0.9; 1.2)	0.43	28%	2	240	1.1 (1.0; 1.3)	0.09	0%
6-month FU	3	360	1.1 (0.9; 1.4)	0.50	72%	2	240	1.0 (0.8; 1.1)	0.66	0%

A negative MD indicates a favour for ESI; a positive MD indicates a favour for placebo.

CI: confidence interval; FU: follow-up; RR: risk ratio

* The limited number of studies did not allow for sensitivity analysis of heterogeneity

Pooling of success data on disability (≥ 50% improvement in ODI scores) favoured none of the interventions for 3- and 6-month follow-up, while 6-week data was lacking (Table 4) (ESM 9). Sensitivity analyses were not feasible due to similarity in methodology between trials.

GRADE quality of evidence: very low (ESM 12).

Success data from non-epidural placebo studies could not be pooled due to varying definitions of success. Data on proportions of treatment success are summarized in ESM 5.

Treatment effect on pain medication use

Acetaminophen use was significantly more reduced after ESI at 6 weeks, but not at 3 weeks [25]. Diclofenac usage was described in one study [27], but the results seemed unrealistic as an inordinate maximum daily intake of 26 tablets was recorded. (Non)opioid analgesic usage was not found to be different between treatments at one month [26], while stronger NSAID and morphine reductions were observed after ESI at 6 weeks [31]. Three other trials all demonstrated equal morphine equivalents during 2-year follow-up [34–36]. For analgesic usage after ESI and non-epidural placebo [23, 30], no significant

differences between treatment groups were found for up to 1-year follow-up. These data are summarized in ESM 6.

Adverse events

Five articles [28, 30, 33, 34, 37] reported absence of any complication due to ESI or placebo during follow-up. Of the remaining 12 trials, one study [32] described a retroperitoneal hematoma after ESI. Several studies [25, 34–36] mentioned periprocedural complications without adverse consequences: dural punctures (1.5% of procedures) [25, 35], intravascular infiltrations (4.1%) [29, 36], nerve root irritations (1.5%) [36] and vasovagal response after placebo (0.8%) [29]. Several minor adverse events similarly occurred in both treatment arms: headache (14.2%) [15, 23, 25, 27, 38], local pain (15.6%) [27, 31], tinnitus (5.5%) [27] and nausea (8.2%) [23, 27]. In the steroid group, single cases of weight gain [27], nonlocal rash [26], irregular periods for several months [24] and two patients with backache and hypotension [15] were reported. In the placebo group, temporary worsening of pain (10.0%) [26] and one case of thoracic pain [38] were described.

DISCUSSION

This review has demonstrated that ESI results in significantly greater leg pain relief and functional improvement

compared to epidural placebo at 6 weeks in patients with sciatica. Caudal and TF injections provided more leg pain relief than IL injections and patients with radiologically confirmed lumbar disc herniation benefitted more than clinically diagnosed patients. In comparison with non-epidural placebo, ESI did not result in more improved leg pain at 6 weeks, although in patients explicitly diagnosed with disc herniation ESI had considerably more effect. For disability and back pain differences were smaller and non-significant.

At 3 months, ESI only resulted in better improvement of leg pain compared to epidural placebo. For other 3-month and all 6-month outcomes, ESI did not demonstrate greater efficacy regardless of the type of placebo. However, these results are to be expected since sciatica is considered a self-limiting condition with a favourable prognosis and steroids are presumed to have only a temporary effect of weeks to months that attenuates gradually. For epidural placebo, TF and IL steroid injections were generally more effective than the caudal approach while for non-epidural placebo differences between epidural routes were less distinct. Differences in treatment efficacy between ESI and placebo injections

were mostly not statistically significant and overall, quality of evidence was moderate to very low.

Several authors consider control injections into the epidural space to be no true placebos, due to their assumed physiological and mechanical effects [39–42]. Instead, they assume that non-epidural injections with an inactive injectate into inactive tissues are genuine placebos. The effect estimate between ESI and non-epidural placebo therapy would expectedly be larger than between ESI and epidural placebo. Interestingly, pooled estimates between ESI and epidural placebo more often favoured steroids than comparisons between ESI and non-epidural placebo. However, due to the very limited number of non-epidural placebo studies it is possible that this result may be changed by conclusions from future trials.

Although several trials have compared ESI to (non-)epidural placebo for sciatica, the study populations and methods broadly differ. Varying aetiology, duration of symptoms, injection contents, placebo types and concomitant therapies introduce clinical heterogeneity which can lead to inaccurate effect estimates and reduced generalizability complicating appropriate conclusions for clinical practice [11, 43, 44]. In order to minimize clinical heterogeneity, all three epidural approaches and placebo types were assessed separately. Additionally, only studies that addressed sciatica patients with a clinical diagnosis or with radiologically confirmed LDH were accepted and assessed independently in sensitivity analyses among other variables. Furthermore, all effect estimates were calculated for three follow-up periods, which allowed for the use of multiple data from primary studies and minimized pooling of less compatible data (e.g. 1-month data used for 3-month effect). Hence, the results in our review are based on analyses of patient groups and treatments with maximized clinical homogeneity.

While ESI induces significantly greater improvement compared to epidural placebo at 6 weeks and 3 months, the absolute treatment differences appear to be modest. A minimally clinical important difference (MCID) for pain and disability has been proposed by a consensus group of experts of, respectively, 10 and 15 points, both on a 0–100 scale [45]. Several of the included studies demonstrated results not meeting this MCID and consequently, the pooled effect estimates are lower than the proposed thresholds. However, the effect of ESI may be obscured by the use of continuous data since 'responders' and 'non-responders' exist [28, 46]. Hence, categorical data based on a pre-defined cut-off condition may be more suitable and is common practice in spinal intervention research, but authors often use variable definitions for 'success' and 'failure'

strongly reducing comparability. Therefore, we only pooled studies with the same definition of 'success'. simultaneously, omitting studies with other definitions may introduce bias. In this review, no significant differences in treatment success based on categorical data were observed. The use of a standardized definition in future trials would allow for pooling of more studies that could affect this result [45, 47–50]. Additionally, identification of subgroups of patients more responsive to ESI could justify steroid injections for these specific groups [28, 35, 51]. Although some sensitivity analyses identified more responsive subgroups (e.g. radiological versus clinical diagnosis for 6-week leg pain), these results must be interpreted with caution due to the limited number of studies included in each analysis.

In addition to treatment efficacy, complications must be considered when reviewing the validity of ESI therapy in clinical practice. Complications can generally be associated with needle placement or with administration of corticosteroids. In our review, only one patient with a serious complication and several minor events were described. This is in accordance with the observations from large cohort studies that the most frequently reported complications are minor and transient, but serious complications can develop, although very rarely and that, with correct safety measures, ESI can be considered a safe therapy [52–58].

Our review is limited by the paucity of literature on ESI for sciatica. A relatively small number of studies was eligible for inclusion particularly fo r comparisons with non-epidural placebo. Therefore, sensitivity analyses for each epidural approach separately were not feasible. Additionally, this paucity of literature complicated the interpretation of sensitivity analyses and publication bias [59]. Furthermore, the wide variety of definitions of treatment success precluded pooling of all available studies for categorical data.

The increasing demand for evidence-based medicine calls for studies with appropriate methodological quality and applicability [60–62]. Future studies that compare ESI to (non-)epidural placebos should consider the aetiology of sciatica and carefully monitor concomitant therapy. Moreover, studies should use a standardized cut-off condition for treatment success and focus on clinical, radiological and pathological features that can differentiate between responders and non-responders.

With the current evidence, ESI can be recommended as short-term pain management therapy compared to epidural placebo, although it must be stressed that generally MCID is not met. ESI has no proven additional value at 3 and 6

months or compared to non-epidural placebo at present. Absolute treatment differences are modest, but possibly subgroups of patients exist that will benefit more than others. With appropriate safety measures, ESI is a safe treatment. In clinical practice, physicians and patients should discuss the possible small short-term benefits and complications of ESI in a process of shared decision-making.

Electronic Supplemental Materials (ESM) can be reviewed online:
https://doi.org/10.1007/s00586-021-06854-9

REFERENCES

1. Konstantinou K, Dunn KM (2008) Sciatica: review of epidemiological studies and prevalence estimates. Spine (Phila Pa 1976) 33 (22):2464-2472. doi:10.1097/BRS.0b013e318183a4a2
2. Koes BW, van Tulder MW, Peul WC (2007) Diagnosis and treatment of sciatica. Bmj 334 (7607):1313-1317. doi:10.1136/bmj.39223.428495.BE
3. Deyo RA, Mirza SK (2016) CLINICAL PRACTICE. Herniated Lumbar Intervertebral Disk. N Engl J Med 374 (18):1763-1772. doi:10.1056/NEJMcp1512658
4. Stafford MA, Peng P, Hill DA (2007) Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth 99 (4):461-473. doi:10.1093/bja/aem238
5. Janardhana AP, Rajagopal, Rao S, Kamath A (2010) Correlation between clinical features and magnetic resonance imaging findings in lumbar disc prolapse. Indian J Orthop 44 (3):263-269. doi:10.4103/0019-5413.65148
6. van Rijn JC, Klemetso N, Reitsma JB, Majoie CB, Hulsmans FJ, Peul WC, Bossuyt PM, Heeten GJ, Stam J (2006) Symptomatic and asymptomatic abnormalities in patients with lumbosacral radicular syndrome: Clinical examination compared with MRI. Clin Neurol Neurosurg 108 (6):553-557. doi:10.1016/j.clineuro.2005.10.003
7. el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, Van der Kallen BF, van den Hout WB, Jacobs WC, Koes BW, Peul WC (2013) Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med 368 (11):999-1007. doi:10.1056/NEJMoa1209250
8. McLain RF, Kapural L, Mekhail NA (2005) Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J 5 (2):191-201. doi:10.1016/j.spinee.2004.10.046
9. Mulleman D, Mammou S, Griffoul I, Watier H, Goupille P (2006) Pathophysiology of disk-related sciatica. I.--Evidence supporting a chemical component. Joint Bone Spine 73 (2):151-158. doi:10.1016/j.jbspin.2005.03.003
10. Wang K, Bao JP, Yang S, Hong X, Liu L, Xie XH, Wu XT (2016) A cohort study comparing the serum levels of pro- or anti-inflammatory cytokines in patients with lumbar radicular pain and healthy subjects. Eur Spine J 25 (5):1428-1434. doi:10.1007/s00586-015-4349-4
11. Valat JP, Genevay S, Marty M, Rozenberg S, Koes B (2010) Sciatica. Best Pract Res Clin Rheumatol 24 (2):241-252. doi:10.1016/j.berh.2009.11.005
12. Knezevic NN, Jovanovic F, Voronov D, Candido KD (2018) Do Corticosteroids Still Have a Place in the Treatment of Chronic Pain? Front Pharmacol 9:1229. doi:10.3389/fphar.2018.01229
13. Collighan N, Gupta S (2009) Epidural steroids. Continuing Education in Anaesthesia Critical Care & Pain 10 (1):1-5. doi:10.1093/bjaceaccp/mkp043
14. Starkweather A, Witek-Janusek L, Mathews HL (2005) Neural-immune interactions: implications for pain management in patients with low-back pain and sciatica. Biol Res Nurs 6 (3):196-206. doi:10.1177/1099800404272221
15. Nandi J, Chowdhery A (2017) A Randomized Controlled Clinical Trial to Determine the Effectiveness of Caudal Epidural Steroid Injection in Lumbosacral Sciatica. J Clin Diagn Res 11 (2):Rc04-rc08. doi:10.7860/jcdr/2017/21905.9392
16. Manchikanti L, Falco FJ, Singh V, Pampati V, Parr AT, Benyamin RM, Fellows B, Hirsch JA (2012) Utilization of interventional techniques in managing chronic pain

in the Medicare population: analysis of growth patterns from 2000 to 2011. Pain Physician 15 (6):E969-982
17. Friedly J, Chan L, Deyo R (2007) Increases in lumbosacral injections in the Medicare population: 1994 to 2001. Spine (Phila Pa 1976) 32 (16):1754-1760. doi:10.1097/BRS.0b013e3180b9f96e
18. Manchikanti L, Pampati V, Falco FJ, Hirsch JA (2013) Assessment of the growth of epidural injections in the medicare population from 2000 to 2011. Pain Physician 16 (4):E349-364
19. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 366:l4898. doi:10.1136/bmj.l4898
20. Higgins JPT, Li T, Deeks JJe (2019) Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J et al. (eds) Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane,
21. Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. Bmj 315 (7109):629-634. doi:10.1136/bmj.315.7109.629
22. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, Bronfort G, van Tulder MW (2015) 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976) 40 (21):1660-1673. doi:10.1097/brs.0000000000001061
23. Arden NK, Price C, Reading I, Stubbing J, Hazelgrove J, Dunne C, Michel M, Rogers P, Cooper C (2005) A multicentre randomized controlled trial of epidural corticosteroid injections for sciatica: the WEST study. Rheumatology (Oxford) 44 (11):1399-1406. doi:10.1093/rheumatology/kei028
24. Bush K, Hillier S (1991) A controlled study of caudal epidural injections of triamcinolone plus procaine for the management of intractable sciatica. Spine (Phila Pa 1976) 16 (5):572-575. doi:10.1097/00007632-199105000-00015
25. Carette S, Leclaire R, Marcoux S, Morin F, Blaise GA, St-Pierre A, Truchon R, Parent F, Levésque J, Bergeron V, Montminy P, Blanchette C (1997) Epidural corticosteroid injections for sciatica due to herniated nucleus pulposus. N Engl J Med 336 (23):1634-1640. doi:10.1056/nejm199706053362303
26. Cohen SP, White RL, Kurihara C, Larkin TM, Chang A, Griffith SR, Gilligan C, Larkin R, Morlando B, Pasquina PF, Yaksh TL, Nguyen C (2012) Epidural steroids, etanercept, or saline in subacute sciatica: a multicenter, randomized trial. Ann Intern Med 156 (8):551-559. doi:10.7326/0003-4819-156-8-201204170-00002
27. Datta R, Upadhyay KK (2011) A Randomized Clinical Trial of Three Different Steroid Agents for Treatment of Low Backache through the Caudal Route. Med J Armed Forces India 67 (1):25-33. doi:10.1016/s0377-1237(11)80007-9
28. Ghahreman A, Ferch R, Bogduk N (2010) The efficacy of transforaminal injection of steroids for the treatment of lumbar radicular pain. Pain Med 11 (8):1149-1168. doi:10.1111/j.1526-4637.2010.00908.x
29. Ghai B, Kumar K, Bansal D, Dhatt SS, Kanukula R, Batra YK (2015) Effectiveness of Parasagittal Interlaminar Epidural Local Anesthetic with or without Steroid in

Chronic Lumbosacral Pain: A Randomized, Double-Blind Clinical Trial. Pain Physician 18 (3):237-248
30. Helliwell M, Robertson J, Ellis R (1985) Outpatient treatment of low-back pain and sciatica by a single extradural corticosteroid injection. British Journal of Clinical Practice 39 (6):228-231
31. Iversen T, Solberg TK, Romner B, Wilsgaard T, Twisk J, Anke A, Nygaard O, Hasvold T, Ingebrigtsen T (2011) Effect of caudal epidural steroid or saline injection in chronic lumbar radiculopathy: multicentre, blinded, randomised controlled trial. Bmj 343:d5278. doi:10.1136/bmj.d5278
32. Karppinen J, Malmivaara A, Kurunlahti M, Kyllönen E, Pienimäki T, Nieminen P, Ohinmaa A, Tervonen O, Vanharanta H (2001) Periradicular infiltration for sciatica: a randomized controlled trial. Spine (Phila Pa 1976) 26 (9):1059-1067. doi:10.1097/00007632-200105010-00015
33. Klenerman L, Greenwood R, Davenport HT, White DC, Peskett S (1984) Lumbar epidural injections in the treatment of sciatica. Br J Rheumatol 23 (1):35-38. doi:10.1093/rheumatology/23.1.35
34. Manchikanti L, Singh V, Cash KA, Pampati V, Damron KS, Boswell MV (2012) Effect of fluoroscopically guided caudal epidural steroid or local anesthetic injections in the treatment of lumbar disc herniation and radiculitis: a randomized, controlled, double blind trial with a two-year follow-up. Pain Physician 15 (4):273-286
35. Manchikanti L, Singh V, Cash KA, Pampati V, Falco FJ (2014) A randomized, double-blind, active-control trial of the effectiveness of lumbar interlaminar epidural injections in disc herniation. Pain Physician 17 (1):E61-74
36. Manchikanti L, Cash KA, Pampati V, Falco FJ (2014) Transforaminal epidural injections in chronic lumbar disc herniation: a randomized, double-blind, active-control trial. Pain Physician 17 (4):E489-501
37. Tafazal S, Ng L, Chaudhary N, Sell P (2009) Corticosteroids in peri-radicular infiltration for radicular pain: a randomised double blind controlled trial. One year results and subgroup analysis. Eur Spine J 18 (8):1220-1225. doi:10.1007/s00586-009-1000-2
38. Valat JP, Giraudeau B, Rozenberg S, Goupille P, Bourgeois P, Micheau-Beaugendre V, Soubrier M, Richard S, Thomas E (2003) Epidural corticosteroid injections for sciatica: a randomised, double blind, controlled clinical trial. Ann Rheum Dis 62 (7):639-643. doi:10.1136/ard.62.7.639
39. Bicket MC, Gupta A, Brown CHt, Cohen SP (2013) Epidural injections for spinal pain: a systematic review and meta-analysis evaluating the "control" injections in randomized controlled trials. Anesthesiology 119 (4):907-931. doi:10.1097/ALN.0b013e31829c2ddd
40. Cohen SP, Bicket MC, Jamison D, Wilkinson I, Rathmell JP (2013) Epidural steroids: a comprehensive, evidence-based review. Reg Anesth Pain Med 38 (3):175-200. doi:10.1097/AAP.0b013e31828ea086
41. Manchikanti L, Knezevic NN, Boswell MV, Kaye AD, Hirsch JA (2016) Epidural Injections for Lumbar Radiculopathy and Spinal Stenosis: A Comparative Systematic Review and Meta-Analysis. Pain Physician 19 (3):E365-410
42. Manchikanti L, Nampiaparampil DE, Manchikanti KN, Falco FJ, Singh V, Benyamin RM, Kaye AD, Sehgal N, Soin A, Simopoulos TT, Bakshi S, Gharibo CG, Gilligan CJ, Hirsch JA (2015) Comparison of the efficacy of saline, local anesthetics, and steroids in epidural and facet joint injections for the management of spinal pain:

A systematic review of randomized controlled trials. Surg Neurol Int 6 (Suppl 4):S194-235. doi:10.4103/2152-7806.156598
43. Deeks JJ, Higgins JPT, Altman DGe (2019) Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J et al. (eds) Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane,
44. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, Sun X, Moher D (2013) Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. BMC Med Res Methodol 13:106. doi:10.1186/1471-2288-13-106
45. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, Bouter LM, de Vet HC (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 33 (1):90-94. doi:10.1097/BRS.0b013e31815e3a10
46. MacVicar J, King W, Landers MH, Bogduk N (2013) The effectiveness of lumbar transforaminal injection of steroids: a comprehensive review with systematic analysis of the published data. Pain Med 14 (1):14-28. doi:10.1111/j.1526-4637.2012.01508.x
47. Haefeli M, Elfering A (2006) Pain assessment. Eur Spine J 15 Suppl 1 (Suppl 1):S17-24. doi:10.1007/s00586-005-1044-x
48. Mehling WE, Gopisetty V, Acree M, Pressman A, Carey T, Goldberg H, Hecht FM, Avins AL (2011) Acute low back pain and primary care: how to define recovery and chronification? Spine (Phila Pa 1976) 36 (26):2316-2323. doi:10.1097/BRS.0b013e31820c01a6
49. Moore A, McQuay H, Gavaghan D (1996) Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. Pain 66 (2-3):229-237. doi:10.1016/0304-3959(96)03032-1
50. Vavken P, Ganal-Antonio AK, Quidde J, Shen FH, Chapman JR, Samartzis D (2015) Fundamentals of Clinical Outcomes Assessment for Spinal Disorders: Clinical Outcome Instruments and Applications. Global Spine J 5 (4):329-338. doi:10.1055/s-0034-1396046
51. Karppinen J, Ohinmaa A, Malmivaara A, Kurunlahti M, Kyllönen E, Pienimäki T, Nieminen P, Tervonen O, Vanharanta H (2001) Cost effectiveness of periradicular infiltration for sciatica: subgroup analysis of a randomized controlled trial. Spine (Phila Pa 1976) 26 (23):2587-2595. doi:10.1097/00007632-200112010-00013
52. Abdi S, Datta S, Lucas LF (2005) Role of epidural steroids in the management of chronic spinal pain: a systematic review of effectiveness and complications. Pain Physician 8 (1):127-143
53. Anesthesiologie SP-ePGvdNVv (2016) Protocol Veilig gebruik van epidurale corticosteroïden infiltraties. https://www.anesthesiologie.nl/publicaties/protocol-veilig-gebruik-epidurale-corticosteroiden-infiltraties.
54. Chou R, Atlas SJ, Stanos SP, Rosenquist RW (2009) Nonsurgical interventional therapies for low back pain: a review of the evidence for an American Pain Society clinical practice guideline. Spine (Phila Pa 1976) 34 (10):1078-1093. doi:10.1097/BRS.0b013e3181a103b1
55. El-Yahchouchi CA, Plastaras CT, Maus TP, Carr CM, McCormick ZL, Geske JR, Smuck M, Pingree MJ, Kennedy DJ (2016) Adverse Event Rates Associated with

Transforaminal and Interlaminar Epidural Steroid Injections: A Multi-Institutional Study. Pain Med 17 (2):239-249. doi:10.1111/pme.12896

56. Manchikanti L, Malla Y, Wargo BW, Cash KA, Pampati V, Fellows B (2012) A prospective evaluation of complications of 10,000 fluoroscopically directed epidural injections. Pain Physician 15 (2):131-140
57. McGrath JM, Schaefer MP, Malkamaki DM (2011) Incidence and characteristics of complications from epidural steroid injections. Pain Med 12 (5):726-731. doi:10.1111/j.1526-4637.2011.01077.x
58. Carr CM, Plastaras CT, Pingree MJ, Smuck M, Maus TP, Geske JR, El-Yahchouchi CA, McCormick ZL, Kennedy DJ (2016) Immediate Adverse Events in Interventional Pain Procedures: A Multi-Institutional Study. Pain Med 17 (12):2155-2161. doi:10.1093/pm/pnw051
59. Debray TPA, Moons KGM, Riley RD (2018) Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests. Res Synth Methods 9 (1):41-50. doi:10.1002/jrsm.1266
60. Pearce W, Raman S (2014) The new randomised controlled trials (RCT) movement in public policy: challenges of epistemic governance. Policy Sciences 47 (4):387-402. doi:10.1007/s11077-014-9208-3
61. Rogers W, Hutchison K (2017) Evidence-Based Medicine in Theory and Practice: Epistemological and Normative Issues. In: Schramme T, Edwards S (eds) Handbook of the Philosophy of Medicine. Springer, Dordrecht,
62. Victora CG, Habicht JP, Bryce J (2004) Evidence-based public health: moving beyond randomized trials. Am J Public Health 94 (3):400-405. doi:10.2105/ajph.94.3.400

3

The outcome of epidural injections in lumbar radiculopathy is not dependent on the presence of disc herniation on magnetic resonance imaging: assessment of short term and long-term efficacy

E.J.A. Verheijen, A.G. Munts, O.B.H.A.M. van Haagen,
D. de Vries, C.L.A. Vleggeert-Lankamp

World Neurosurg 2021

ABSTRACT

Objective

Lumbar radiculopathy is a condition with major physical, social, and economic consequences. Despite its favorable prognosis, the burden can be significant. In this study, we aimed to determine the value of magnetic resonance imaging (MRI) and the efficacy of transforaminal epidural injections (TEIs) in patients with lumbar radiculopathy secondary to lumbar disc herniation (LDH) and other causes (non-LDH).

Methods

Patients with lumbar radiculopathy were reviewed for radiologic diagnosis based on MRI. For patients receiving TEI therapy, response after 6–8 weeks (short-term) and 16 weeks (long-term), number of injections, subsequent surgery, and patient outcome were evaluated. Treatment response was assessed by patient-reported symptom relief and numeric rating scale pain scores.

Results

Overall, 66% of MRI examinations showed a clinically relevant LDH. A total of 486 of 1824 patients received TEI, of whom one third did not show LDH. Of patients, 70% reported a short-term effect with significant pain reduction and 44% reported a long-term effect. No significant differences were observed between the LDH and non-LDH groups. Of patients, 59% required multiple injections and reported similar efficacy compared with patients treated with a single injection.

Conclusions

A considerable part of MRI examinations in patients with lumbar radiculopathy do not show a clinically relevant LDH. Regardless of the radiologic diagnosis, most patients treated with TEI benefit in both the short-term and the long-term after a single-injection or multiple-injection regime. Subsequent injections are advisable if the effect from the first injection is unsatisfactory or wears off. MRI examination before TEI therapy may be redundant, which allows for expedition of this treatment.

INTRODUCTION

Lumbar radiculopathy is a common spinal condition characterized by radicular pain toward the lower extremity.[1] Although lumbar disc herniation (LDH) is the most prevalent cause, a variety of degenerative spinal disorders can instigate unilateral radicular symptoms originating from the lumbar spine.[2] It is assumed that the origin of lumbar radiculopathy is multifactorial, involving a complex interplay of physical impingement, inflammatory processes, and immunologic responses resulting from degenerative spinal changes, which, consequently, induce lumbosacral nerve root irritation.[3,4] Patients often present with severe leg pain, paresthesia, and other neurologic deficits that seriously impede their daily functioning and decrease their quality of life.[5]

Despite the severity of symptoms, most patients with lumbar radiculopathy spontaneously recover within 1 year.[6,7] Therefore, patients are initially treated conservatively for several weeks before consultation with a neurologist and extensive radiologic imaging,[8] considering the absence of warning signs that require immediate attention. After referral, surgical options can be discussed with the neurosurgeon, but generally, patients are encouraged to prolong conservative therapy up to 16 weeks.[9] Nevertheless, the physical and mental burden while awaiting natural resolution can be intolerable.

Transforaminal epidural injection (TEI) therapy is a widely used treatment to create a more bearable situation for these patients.[10,11] A small amount of local anesthetic and corticosteroid is administered in the vicinity of the affected nerve root to alleviate symptoms by reducing the surrounding inflammatory processes.[12] Treatment with TEI is intended as a temporary pain management strategy because the effect of a single injection may wear off after a couple of weeks to months. In part of the patients, this treatment is sufficient to await natural resolution of symptoms, obviating further treatment.[13]

Although TEI has been a long-established treatment for lumbar radiculopathy, controversy remains over its efficacy and the correlation with radiologic findings because of discrepancies between studies and a paucity of literature for causes other than LDH.[8,]14, 15, 16, 17, 18, 19, 20, 21 Moreover, it is still unclear whether multiple injections can produce equal long-term improvement compared with a single injection. In addition, it has been hypothesized that short-term response to TEI might predict long-term patient outcome, making it useful as a clinical decision-making tool.22, 23, 24

The main goal of this study is to retrospectively assess the correlation between the presence of a herniated disc on magnetic resonance imaging (MRI) and outcome of TEI in patients with lumbar radiculopathy in clinical practice. Second, the effect of a regimen of multiple injections, difference in surgery frequency, and patient outcome after treatment are evaluated.

METHODS

Patients

For this large retrospective cohort study, data from the Spaarne Gasthuis, Hoofddorp/Haarlem, the Netherlands, were used with consent from the medical ethical committee and the board of directors. The hospital's electronic patient record system was searched to select all patients who were diagnosed with lumbar radiculopathy between January 1, 2016 and September 1, 2017. The following inclusion criteria were used: 1) clinical diagnosis of lumbar radiculopathy by a neurologist or neurosurgeon, 2) radiologic examination of the lumbar and sacral spine with MRI, and 3) patient age of 18 years or older.

Data collection

For all patients data was collected concerning MRI examination and primary treatment choice. For patients treated with TEI, additional data were collected regarding demographics, cause of symptoms as deduced from MRI examination by the neurologist or neurosurgeon, history of back surgery, analgesic use, level of primary injection, additional injections, surgical intervention, and patient outcome after treatment. Response to TEI in the short-term was evaluated after 6–8 weeks after the first epidural injection by the anesthesiologist, neurologist, or nurse by telephone. Patients reported overall pain scores (numeric rating scale [NRS] ranging from 0 [no pain] to 10 [worst imaginable pain]) and whether they considered the treatment to be effective. Long-term response to TEI was evaluated at 16 weeks by patient-reported symptom relief. If patients were subjected to spinal surgery subsequent to TEI within these 16 weeks, TEI therapy was considered to have failed for the long-term. Patient outcome at the end of treatment was evaluated at the last hospital appointment for this episode of sciatic symptoms using a Likert scale with 3 categories, defined by "unsatisfactory", "satisfactory," and "good" outcome. Unsatisfactory was defined as no or only slight improvement compared with the patient's initial situation. A satisfactory outcome constituted significant improvement, but with the patient still experiencing some degree of pain or disability. A good outcome signified nearly complete or full recovery.

Treatment

Patients visited the outpatient clinic of the anesthesiology department to receive treatment with a TEI. The procedures were performed by several experienced anesthesiologists. The segmental level to be treated was chosen by the anesthesiologist based on MRI results and clinical findings. During the TEI procedure, the patient was positioned in a prone position on the table. After verification of the correct side and level, the skin was disinfected. Under fluoroscopic guidance with anteroposterior and lateral views, a spinal needle was inserted into the targeted neuroforamen. To confirm correct positioning of the needle, a contrast agent was injected, showing a neurogram. If the needle was mistakenly placed, it was repositioned and verified for epidural placement into the neuroforamen. A combination of a local anesthetic (lidocaine, chirocaine, or bupivacaine) and a corticosteroid (methylprednisolone 20–80 mg or dexamethasone 7.5–20 mg) was administered. The procedure lasted 15 minutes and afterward, the patient was monitored for half an hour in the recovery room by the nurse. If the first injection yielded an unsatisfactory outcome, a second injection was offered to the patient, either at the same or an adjacent level. Moreover, if the first injection was satisfactory but the effect gradually decreased, subsequent injections were offered. The injection regimen was continued until symptoms had sufficiently ameliorated, subsequent injections were unsuccessful to relieve symptoms, or the patient had decided to discontinue treatment. Patients were offered surgery by the neurosurgeon if conservative treatment failed and MRI showed a clinically relevant operable component.

Analysis

Patients were stratified according to findings derived from radiologic imaging. One group included patients with MRI-confirmed LDH concordant with clinical findings. Disc herniation was defined as a radiologically evident bulging disc, protrusion, or extrusion of an intervertebral disc in the lumbosacral spine. The non-LDH group comprised patients with LDH inconsistent with clinical findings, a degenerative cause of lumbar radiculopathy, or MRI examination without an evident cause. Consequently, patients were categorized based on primary treatment choice. For patients initially treated with TEI, treatment efficacy was extensively evaluated. Statistical analysis was performed using SPSS version 25 (IBM Corp., Armonk, New York, USA). Missing values were imputed adjusting for covariates. Continuous data were presented using mean and standard deviation and compared between groups with the unpaired *t* test and within groups with the paired *t* test. Percentages and exact numbers were used to present categorical data and the χ^2 test was used for comparisons. A *P* value <0.05 was considered statistically significant.

RESULTS

Patient population

A total of 1824 patients were diagnosed with lumbar radiculopathy and subjected to an MRI of the lumbar spine. In 66% (*n* = 1200), LDH concordant with the clinical condition was observed. In 53% (*n* = 630) of these patients with LDH, conservative treatment was prolonged, 19% (*n* = 233) underwent surgery, and 28% (*n* = 337) were referred to the anesthesiologist for TEI. Of 624 patients without radiologic evidence of a clinically relevant herniated disc, 69% (*n* = 430) continued conservative therapy, 7% (*n* = 45) were surgically treated, and 24% (*n* = 149) received TEI therapy.

A total of 486 patients were treated with 1 or multiple epidural injections at the anesthesiology outpatient clinic: 337 patients with a clinically important LDH and 149 patients without LDH. Patients received an average of 1.7 injections until 16 weeks follow-up. Demographic and clinical characteristics are shown in Table 1. Patients with LDH were on average 8.21 (95% confidence interval, 5.31–11.12) years younger, had less often chronic symptoms, and had less often been subjected to previous back surgery. Females were more represented in the non-LDH group. Baseline NRS for overall pain in the LDH group was comparable to the non-LDH group (0.2; 95% confidence interval, −0.6 to 0.3). Use of analgesic medication was equal between the 2 groups. However, when the type of analgesic was considered, patients with LDH-induced radiculopathy more frequently used nonsteroidal antiinflammatory drugs and opioids, although this latter type of medication was prescribed for a significant number of patients in both groups.

Table 1 Baseline characteristics

	LDH group (n = 337)	Non-LDH group (n = 149)	P-value
Age, yrs. (mean ± SD)	55.12 ± 15.28	63.34 ± 14.42	0.000
Sex: M/F	162 (48.1%) / 175 (51.9%)	51 (34.2%) / 98 (65.8%)	0.005
Duration of symptoms			0.006
≤ 3 months	91 (27.0%)	23 (15.4%)	
> 3 months	246 (73.0%)	126 (84.6%)	
Radiological diagnosis			0.000
LDH	337 (100.0%)	0 (0.0%)	
Degenerative stenosis	0 (0.0%)	97 (65.1%)	
Cyst	0 (0.0%)	11 (7.4%)	
Scar tissue	0 (0.0%)	11 (7.4%)	
Spondylolisthesis	0 (0.0%)	5 (3.4%)	

Table 1 Baseline characteristics (*continued*)

	LDH group (n = 337)	Non-LDH group (n = 149)	P-value
Inconclusive	0 (0.0%)	25 (16.8%)	
History of back surgery	63 (18.7%)	41 (27.5%)	0.029
Same level	47 (74.6%)	30 (73.2%)	0.798
Other level	16 (25.3%)	11 (26.8%)	
Level of injection			0.112
L1	1 (0.3%)	2 (1.3%)	
L2	4 (1.2%)	2 (1.3%)	
L3	27 (8.0%)	12 (8.1%)	
L4	52 (15.4%)	25 (16.8%)	
L5	167 (49.6%)	85 (57.0%)	
L6	3 (0.9%)	2 (1.3%)	
S1	83 (24.6%)	20 (13.4%)	
S3	0 (0.0%)	1 (0.7%)	
NRS overall pain (mean ± SD)	8.0 ± 1.4	7.8 ± 1.4	0.471
Pain medication use	307 (91.1%)	138 (92.6%)	0.511
Paracetamol	210 (62.3%)	102 (68.4%)	0.212
NSAIDs	168 (49.9%)	53 (35.6%)	0.003
Opioids	202 (59.9%)	75 (50.3%)	0.042
Antineuropathic drugs (gabapentin, pregabalin, amitriptyline)	71 (21.1%)	23 (15.4%)	0.157

Short-term efficacy of TEI

Of the 337 patients with LDH receiving TEI, 71% (*n* = 239) reported pain relief between 6 and 8 weeks after the first injection, whereas in the non-LDH group, 68% (*n* = 102) experienced a pain-reducing effect after treatment with an epidural injection. The mean NRS score decreased significantly after treatment to 4.7 ± 2.8 and 4.5 ± 2.6 for, respectively, the LDH group and non-LDH group compared with baseline. However, the difference in NRS reduction between groups was not statistically significant (*P* = 0.868) (Table 2).

Table 2 Outcome parameters at follow-up after TEI therapy between LDH and non-LDH patients

	LDH group (n = 337)	Non-LDH group (n = 149)	P-value
Short-term effect (y/n)	239 (71%) / 98 (29%)	102 (68%) / 47 (32%)	0.491
NRS overall pain (mean ± SD)			
At 8-week follow-up	4.7 ± 2.8	4.5 ± 2.6	0.469
Absolute change*	-3.2 ± 3.0	-3.3 ± 2.9	Intragroup: (<0.001 / <0.001) Intergroup: 0.868
Additional injections (y/n)	197 (58%) / 140 (42%)	90 (60%) / 59 (40%)	0.688
Long-term effect (y/n)	149 (44%) / 188 (56%)	63 (42%) / 86 (58%)	0.781
Surgery (y/n)	73 (22%) / 265 (79%)	18 (12%) / 131 (88%)	0.013
Likert treatment outcome (g/s/u)†	58 (17%) / 162 (48%) / 117 (35%)	13 (9%) / 72 (48%) / 64 (43%)	0.029

* Compared to NRS score at baseline

† g = good; s = satisfactory; u = unsatisfactory

Additional injections

Although 71% of patients in the LDH group experienced a relevant pain reduction after TEI, 59% of these patients (*n* = 140) required additional injections. Of patients in the LDH group who did not benefit from the first injection, 58% (*n* = 57) also received additional epidural injections (Figure 1). Similar results were found in the non-LDH group: 61% (*n* = 62) of patients responsive to the first TEI received subsequent injections, whereas this was true for 60% (*n* = 28) of patients unresponsive to the first epidural injection (Figure 2). Hence, a total of 58% (LDH) and 60% (non-LDH) received supplemental epidural injections, which was a nonsignificant difference (*P* = 0.688) (Table 2).

Long-term efficacy, surgery, and patient outcome at end of treatment

Long-term efficacy was assessed at 16 weeks follow-up (range, 13–18 weeks). Ten percent of patients in the LDH group (*n* = 33) and 6% in the non-LDH group (*n* = 8) underwent surgery before the 16-week follow-up moment. In these patients, TEI therapy was considered to have failed in the long-term (Figures 1 and 2). Of patients with LDH, 44% (*n* = 149) compared with 42% (*n* = 63) of patients without LDH reported a positive effect of TEI at 16 weeks (*P* = 0.203). More patients in the LDH group opted for surgery (*P* = 0.013). The outcome at the end of treatment using a 3-point Likert scale showed that a satisfactory or good outcome was more frequently observed in the LDH group (*P* = 0.029) (Table 2).

Patients with long-term effect after one or more epidural injections

Of the 337 patients with LDH, 239 (71%) initially experienced pain relief after the first injection in the short-term. In 24% of those patients (*n* = 57), this effect was still present at 16 weeks if treatment was restricted to a single injection. If additional injections were administered, another 29% (*n* = 70) reported improved symptoms in the long-term. Of the 98 patients with LDH not responding to the first injection, 22% (*n* = 22) indicated improved symptoms at 16 weeks after administration of additional injections. Therefore, 44% of patients with LDH (*n* = 149) showed symptom relief with a regimen of ≥1 injections at 16 weeks follow-up (Figure 1).

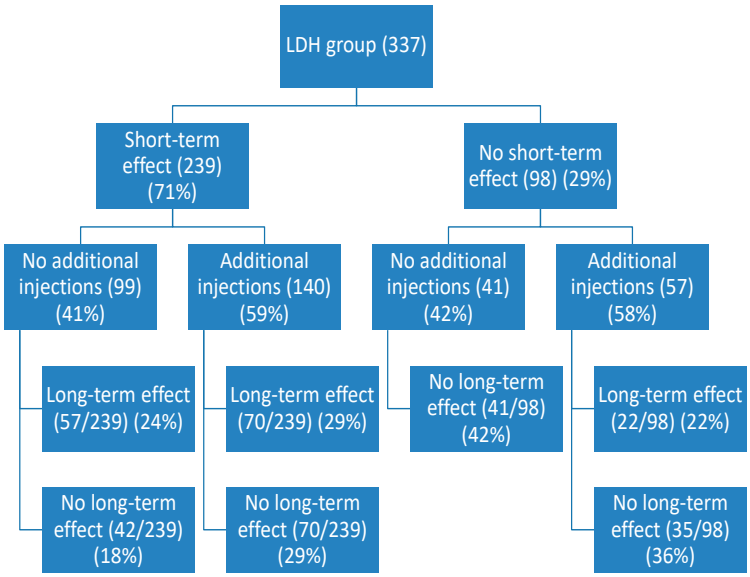


Fig. 1 Summary of transforaminal epidural injection results for patients with lumbar disc herniation (LDH).

Comparable results were observed in the group of 149 patients without LDH. A total of 102 patients (68%) initially experienced a pain-reducing effect after the first injection. If treatment was limited to that single injection, 23% (*n* = 23) reported a long-term effect. However, if patients required additional epidural injections despite a positive effect after the first injection, an additional 29% (*n* = 30) experienced an effect at 16 weeks. In the group of patients without LDH who lacked a response to the initial injection but were subjected to subsequent injections, 22% (*n* = 10) reported a long-term effect. Therefore, for all patients without LDH, 42% (*n* = 63) had improved symptoms at 16 weeks with ≥1 epidural injections (Figure 2).

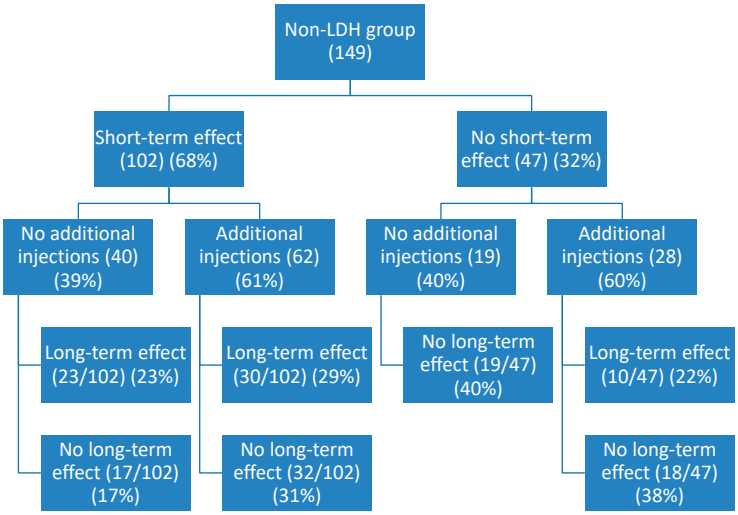


Fig. 2 Summary of transforaminal epidural injection results for patients without lumbar disc herniation (non-LDH).

DISCUSSION

The current study highlights that the efficacy of TEI in patients with lumbar radiculopathy is independent of the radiologic diagnosis. Comparable success rates and pain score reductions were observed in the LDH and non-LDH group for both short-term and long-term follow-up. These results indicate that nerve root compression secondary to a clinically concordant LDH is not a prerequisite for effective treatment with TEI. In the absence of nerve root compression or detectable disc herniation on MRI, inflammatory processes resulting from degenerative changes in other spinal structures may play a fundamental role.25, 26, 27, 28, 29, 30, 31 TEI therapy is assumed to target the inflammatory cascade directly and indirectly through multiple pathways, causing a decrease in nerve root inflammation, nociception of afferent fibers, and activity of proinflammatory cytokines.[3,]32, 33, 34, 35 Therefore, treatment with TEI can be effective in these patients as shown in our study for the non-LDH group.

Controversy exists over symptom improvement because of a placebo effect after TEI, as with other pain management strategies. An injection into the epidural space may achieve a washout of proinflammatory cytokines or result in lysis of inflammatory-mediated adhesions in some patients because injections with only a local anesthetic or saline have been shown to induce some degree of improvement.36, 37, 38 However, several studies have shown transforaminal

steroid injections to be superior compared with placebo in the short-term, vindicating the addition of steroid for TEI.[16,36,39]

The literature on TEI for lumbar spinal conditions other than LDH is limited.[40] A study by Ng et al.[41] found no significant difference in pain reduction between patients with lumbar radiculopathy with LDH and patients with spinal stenosis at 6 weeks follow-up. In addition, Cohen et al.[42] reported that routine MRI before epidural injection had no additional value because equal efficacy after treatment was observed in 127 MRI-blinded and nonblinded patients with lumbar radiculopathy and physicians. These results are in line with results from our study.

A meta-analysis of studies comparing epidural steroid injections with conservative treatment in patients with lumbar radiculopathy[43] found that pain reduction was significantly greater after epidural steroid injections in the short-term. Because the cause of lumbar radiculopathy cannot be correlated with outcome results after TEI, pretreatment MRI examination may not always be necessary. Therefore, TEI therapy could be expedited with extensive imaging postponed to a later stage. Evidently, red flag indications need to be excluded by the general practitioner, concurring with local and national guidelines for treatment of lumbar radiculopathy and referring accordingly. Such a treatment strategy would allow for more adequate symptom management in an earlier stage of lumbar radiculopathy.

A previous study that compared early MRI or computed tomography with delayed imaging in patients with low back pain[44] reported no large effect on overall treatment results. Another trial[45] compared early examination of the lumbar spine with MRI with plain radiography in similar patients and found no significant benefit of MRI scans over radiographic imaging but stressed the higher financial costs of MRI. Hence, plain radiography may be sufficient to detect extreme abnormalities of the lumbar spine that can complicate TEI procedures at an early stage, and periprocedural safety in the absence of extensive imaging can be maintained with, for example, the use of fluoroscopic guidance. Expedition of TEI therapy could lead to improved maintenance of functionality, prevention of financial costs because of absenteeism, and reduction of oral analgesic intake, specifically of opioids, which are prescribed with rapidly increasing frequency despite the serious risk of drug dependence.[46]

Overall, 70% of patients in our study experienced short-term pain relief after a single TEI, with substantial pain reduction. These results are in line with related, although smaller, studies in the medical literature. In a report by Nandi

et al.,[47] 47 patients with LDH received epidural steroid injections, of whom 68% were considered to show success at 4 weeks follow-up. Joswig et al.[23] studied 57 patients with LDH who received TEI, of whom 66.7% were identified as responders after 1 month.

Furthermore, 44% of patients reported long-term pain relief in the current study. Equal long-term effect size could be achieved after treatment with a regimen of multiple injections compared with a single injection in our study population. Additional TEIs were administered to 59% of patients, regardless of their initial response and were beneficial for most patients, resulting in 42%–44% of patients with a satisfactory long-term effect. Moreover, 22% of patients in the LDH group and non-LDH group experienced a good long-term effect after multiple TEIs despite an unsuccessful first injection.

Patients who required additional injections to have a satisfactory outcome at 16 weeks may not have received the primary injection at the appropriate lumbar level. Because the second injection was applied at the adjacent level in a few of these patients, identification of the correct nerve root can be ambiguous at first despite extensive radiologic imaging. Murthy et al.[48] investigated patients with radicular pain and inferred that repeat injections could have a similar effect to the primary injection. Furthermore, when administered within 3 months from the first injection, multiple injections could have a cumulative effect because pain reduction after a subsequent injection was larger than after previous injection(s).

In more recent years, it has been theorized that short-term response to TEI may be used as a predictor for the course of pain during follow-up.22, 23, 24 The rationale is that if the short-term response has predictive value for the severity of symptoms at 16 weeks (the time at which patients are eligible for surgical intervention), it can be used to predict the need for surgery. Early prediction results in surgery at an earlier stage after onset of symptoms. Because of the limited availability of data, it was not feasible to accurately determine whether prediction of long-term patient outcome based on short-term response is possible and this concept should be assessed in a prospective study.

The current study has a few limitations, including its retrospective nature. First, the type and amount of injection contents and exact moment of long-term follow-up varied among patients. Second, only NRS was systematically reported in the patients who received TEI. For all patients who were diagnosed with lumbar radiculopathy and subjected to an MRI of the lumbar spine, it would have been optimal if the clinical condition had been reported using the International

Consortium for Health Outcomes Measurement–dictated outcome measures (i.e., NRS, Oswestry Disability Index, and Quality of Life score [EQ-5D]). The 2 cohorts were not completely similar at baseline. The slightly higher age in the non-LDH group may be explained by the correlation between increasing age and a more degenerative state of the spine.49, 50, 51, 52

The findings of this study indicate that TEI is an effective treatment method to temporarily relieve symptoms for many patients with lumbar radiculopathy regardless of their radiologic diagnosis. MRI examination before TEI may not be necessary and TEI might be planned at an earlier stage, resulting in more adequate treatment of symptoms, reduction of opioid use, and avoidance of surgical intervention. For patients unresponsive or with recurring pain after a single epidural injection, multiple injections could be beneficial. To properly assess the efficacy of TEI in the acute phase of lumbar radiculopathy without extensive radiologic examination and determine the potential predictive value of epidural injections, a randomized controlled trial is required. Moreover, adverse events need to be registered to determine the safety of TEI without preceding MRI.

REFERENCES

1. Schoenfeld AJ, Laughlin M, Bader JO, Bono CM. Characterization of the incidence and risk factors for the development of lumbar radiculopathy. J Spinal Disord Tech. 2012;25(3):163-7.
2. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. Bmj. 2007;334(7607):1313-7.
3. Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth. 2007;99(4):461-73.
4. Ropper AH, Zafonte RD. Sciatica. N Engl J Med. 2015;372(13):1240-8.
5. Hicks GE, Gaines JM, Shardell M, Simonsick EM. Associations of back and leg pain with health status and functional capacity of older adults: findings from the retirement community back pain study. Arthritis Rheum. 2008;59(9):1306-13.
6. Legrand E, Bouvard B, Audran M, Fournier D, Valat JP. Sciatica from disk herniation: Medical treatment or surgery? Joint Bone Spine. 2007;74(6):530-5.
7. el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, Van der Kallen BF, van den Hout WB, Jacobs WC, et al. Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med. 2013;368(11):999-1007.
8. Schaafstra A, Spinnewijn W, Bons S, Borg M, Koes BW, Ostel R, et al. NHG-Standaard Lumbosacraal radiculair syndroom (Tweede herziening). Huisarts Wet. 2015;85(6):308-20.
9. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, et al. Surgery versus prolonged conservative treatment for sciatica. N Engl J Med. 2007;356(22):2245-56.
10. Grøvle L, Haugen AJ, Ihlebaek CM, Keller A, Natvig B, Brox JI, et al. Comorbid subjective health complaints in patients with sciatica: a prospective study including comparison with the general population. J Psychosom Res. 2011;70(6):548-56.
11. Maslak JP, Jenkins TJ, Weiner JA, Kannan AS, Patoli DM, McCarthy MH, et al. Burden of Sciatica on US Medicare Recipients. J Am Acad Orthop Surg. 2020;28(10):e433-e9.
12. Ter Meulen BC, Weinstein H, Ostelo R, Koehler PJ. The Epidural Treatment of Sciatica: Its Origin and Evolution. Eur Neurol. 2016;75(1-2):58-64.
13. Lewis RA, Williams NH, Sutton AJ, Burton K, Din NU, Matar HE, et al. Comparative clinical effectiveness of management strategies for sciatica: systematic review and network meta-analyses. Spine J. 2015;15(6):1461-77.
14. Vroomen PC, de Krom MC, Slofstra PD, Knottnerus JA. Conservative treatment of sciatica: a systematic review. J Spinal Disord. 2000;13(6):463-9.
15. Luijsterburg PA, Verhagen AP, Ostelo RW, van Os TA, Peul WC, Koes BW. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: a systematic review. Eur Spine J. 2007;16(7):881-99.
16. Pinto RZ, Maher CG, Ferreira ML, Hancock M, Oliveira VC, McLachlan AJ, et al. Epidural corticosteroid injections in the management of sciatica: a systematic review and meta-analysis. Ann Intern Med. 2012;157(12):865-77.
17. Smith CC, McCormick ZL, Mattie R, MacVicar J, Duszynski B, Stojanovic MP. The Effectiveness of Lumbar Transforaminal Injection of Steroid for the Treatment of Radicular Pain: A Comprehensive Review of the Published Data. Pain Med. 2020;21(3):472-87.

18. Bhatia A, Flamer D, Shah PS, Cohen SP. Transforaminal Epidural Steroid Injections for Treating Lumbosacral Radicular Pain from Herniated Intervertebral Discs: A Systematic Review and Meta-Analysis. Anesth Analg. 2016;122(3):857-70.
19. Yuce I, Kahyaoglu O, Ataseven M, Cavusoglu H, Aydin Y. Diagnosis and Treatment of Transforaminal Epidural Steroid Injection in Lumbar Spinal Stenosis. Sisli Etfal Hastan Tip Bul. 2020;54(3):327-32.
20. Chang MC, Lee DG. Outcome of Transforaminal Epidural Steroid Injection According to the Severity of Lumbar Foraminal Spinal Stenosis. Pain Physician. 2018;21(1):67-72.
21. Maus TP, El-Yahchouchi CA, Geske JR, Carter RE, Kaufmann TJ, Wald JT, et al. Imaging Determinants of Clinical Effectiveness of Lumbar Transforaminal Epidural Steroid Injections. Pain Med. 2016;17(12):2176-84.
22. El-Yahchouchi C, Wald J, Brault J, Geske J, Hagen C, Murthy N, et al. Lumbar transforaminal epidural steroid injections: does immediate post-procedure pain response predict longer term effectiveness? Pain Med. 2014;15(6):921-8.
23. Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN. The Value of Short-Term Pain Relief in Predicting the One-Month Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg. 2016;96:323-33.
24. Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN. The Value of Short-Term Pain Relief in Predicting the Long-Term Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg. 2017;107:764-71.
25. Heymans M, Eekhout I. Applied Missing Data Analysis With SPSS and (R)Studio, 1st edn Amsterdam2019 [Available from: https://bookdown.org/mwheymans/bookmi/.
26. Igarashi A, Kikuchi S, Konno S, Olmarker K. Inflammatory cytokines released from the facet joint tissue in degenerative lumbar spinal disorders. Spine (Phila Pa 1976). 2004;29(19):2091-5.
27. Valat JP, Genevay S, Marty M, Rozenberg S, Koes B. Sciatica. Best Pract Res Clin Rheumatol. 2010;24(2):241-52.
28. Dower A, Davies MA, Ghahreman A. Pathologic Basis of Lumbar Radicular Pain. World Neurosurg. 2019;128:114-21.
29. Jungen MJ, Ter Meulen BC, van Osch T, Weinstein HC, Ostelo R. Inflammatory biomarkers in patients with sciatica: a systematic review. BMC Musculoskelet Disord. 2019;20(1):156.
30. Löhr M, Hampl JA, Lee JY, Ernestus RI, Deckert M, Stenzel W. Hypertrophy of the lumbar ligamentum flavum is associated with inflammation-related TGF-β expression. Acta Neurochir (Wien). 2011;153(1):134-41.
31. Choi YK. Lumbar foraminal neuropathy: an update on non-surgical management. Korean J Pain. 2019;32(3):147-59.
32. Lipetz JS. Pathophysiology of inflammatory, degenerative, and compressive radiculopathies. Phys Med Rehabil Clin N Am. 2002;13(3):439-49.
33. Collighan N, Gupta S. Epidural steroids. Continuing Education in Anaesthesia Critical Care & Pain. 2009;10(1):1-5.
34. Lakshmi R, Aravindaswami P. A Review of Etiology Pathogenesis, Treatment of Sciatica. IJSR. 2018;7(1):1477-9.
35. Johansson A, Hao J, Sjölund B. Local corticosteroid application blocks transmission in normal nociceptive C-fibres. Acta Anaesthesiol Scand. 1990;34(5):335-8.

36. Li JY, Xie W, Strong JA, Guo QL, Zhang JM. Mechanical hypersensitivity, sympathetic sprouting, and glial activation are attenuated by local injection of corticosteroid near the lumbar ganglion in a rat model of neuropathic pain. Reg Anesth Pain Med. 2011;36(1):56-62.
37. Ghahreman A, Ferch R, Bogduk N. The efficacy of transforaminal injection of steroids for the treatment of lumbar radicular pain. Pain Med. 2010;11(8):1149-68.
38. Valat JP, Giraudeau B, Rozenberg S, Goupille P, Bourgeois P, Micheau-Beaugendre V, et al. Epidural corticosteroid injections for sciatica: a randomised, double blind, controlled clinical trial. Ann Rheum Dis. 2003;62(7):639-43.
39. Bartynski WS, Jennings RB, Rothfus WE, Agarwal V. Immediate pain response to interlaminar lumbar epidural steroid administration: response characteristics and effects of anesthetic concentration. AJNR Am J Neuroradiol. 2013;34(1):239-46.
40. Cohen SP, White RL, Kurihara C, Larkin TM, Chang A, Griffith SR, et al. Epidural steroids, etanercept, or saline in subacute sciatica: a multicenter, randomized trial. Ann Intern Med. 2012;156(8):551-9.
41. MacVicar J, King W, Landers MH, Bogduk N. The effectiveness of lumbar transforaminal injection of steroids: a comprehensive review with systematic analysis of the published data. Pain Med. 2013;14(1):14-28.
42. Ng LC, Sell P. Outcomes of a prospective cohort study on peri-radicular infiltration for radicular pain in patients with lumbar disc herniation and spinal stenosis. Eur Spine J. 2004;13(4):325-9.
43. Cohen SP, Gupta A, Strassels SA, Christo PJ, Erdek MA, Griffith SR, et al. Effect of MRI on treatment results or decision making in patients with lumbosacral radiculopathy referred for epidural steroid injections: a multicenter, randomized controlled trial. Arch Intern Med. 2012;172(2):134-42.
44. Yang S, Kim W, Kong HH, Do KH, Choi KH. Epidural steroid injection versus conservative treatment for patients with lumbosacral radicular pain: A meta-analysis of randomized controlled trials. Medicine (Baltimore). 2020;99(30):e21283.
45. Gilbert FJ, Grant AM, Gillan MG, Vale LD, Campbell MK, Scott NW, et al. Low back pain: influence of early MR imaging or CT on treatment and outcome--multicenter randomized trial. Radiology. 2004;231(2):343-51.
46. Jarvik JG, Hollingworth W, Martin B, Emerson SS, Gray DT, Overman S, et al. Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. Jama. 2003;289(21):2810-8.
47. National Academies of Sciences E, and Medicine; Health and Medicine Division; Board on Health Sciences Policy; Committee on Pain Management and Regulatory Strategies to Address Prescription Opioid Abuse. Pain Management and the Opioid Epidemic: Balancing Societal and Individual Benefits and Risks of Prescription Opioid Use. Washington (DC): National Academies Press (US); 2017.
48. Nandi J, Chowdhery A. A Randomized Controlled Clinical Trial to Determine the Effectiveness of Caudal Epidural Steroid Injection in Lumbosacral Sciatica. J Clin Diagn Res. 2017;11(2):Rc04-rc8.
49. Murthy NS, Geske JR, Shelerud RA, Wald JT, Diehn FE, Thielen KR, et al. The effectiveness of repeat lumbar transforaminal epidural steroid injections. Pain Med. 2014;15(10):1686-94.
50. Benoist M. Natural history of the aging spine. Eur Spine J. 2003;12 Suppl 2(Suppl 2):S86-9.

51. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage. 2012;20(10):1103-8.
52. Saleem S, Aslam HM, Rehmani MA, Raees A, Alvi AA, Ashraf J. Lumbar disc degenerative disease: disc degeneration symptoms and magnetic resonance image findings. Asian Spine J. 2013;7(4):322-34.
53. Yabuki S, Fukumori N, Takegami M, Onishi Y, Otani K, Sekiguchi M, et al. Prevalence of lumbar spinal stenosis, using the diagnostic support tool, and correlated factors in Japan: a population-based study. J Orthop Sci. 2013;18(6):893-900.

4

Factors associated with patient outcome after transforaminal epidural injection for lumbar disc herniation and stenosis: a systematic review of current literature

E.J.A. Verheijen, N.R.E. van der Vlist, C.L.A. Vleggeert-Lankamp

Submitted

ABSTRACT

Purpose

Lumbar radicular pain is usually caused by a herniated disc (LDH) or spinal stenosis which instigates inflammatory and immunological responses. Transforaminal epidural injections (TEI) with steroids are aimed at reducing these responses in order to relieve pain symptoms and improve physical functionality. However, it is unknown which factors can differentiate between responders and non-responders. The purpose of this systematic review is to elucidate which demographical, clinical and radiological factors are associated with patient outcome after TEI.

Methods

The PubMed, Embase, Web of Science, Cochrane Library, Emcare and Academic Search Premier databases were searched for studies assessing the association between demographic, clinical and radiological variables, and clinical outcome after a unilateral, single-level TEI in patients with lumbar radiculopathy secondary to disc herniation or degenerative stenosis. Risk of bias was evaluated using an adjusted version of the Cowley risk of bias criteria. Outcome measures included patient-reported symptom relief, leg pain or physical functionality scores. A qualitative analysis of included studies was performed.

Results

A total of 36 studies were included assessing 37 different factors and their association with TEI outcome. Among demographic variables there was some evidence for an association between shorter duration of symptoms and superior outcome after TEI. For clinical factors there was limited evidence associating a positive Slump test with better patient outcome. Among radiological factors limited evidence was found for an association with stenosis rather than LDH, more centrally located LDH compared to more laterally locations of the herniation, the absence of transitional vertebrae and a higher degree of nerve root compression with a favorable outcome after TEI.

Conclusion

Several factors were identified that associated with TEI outcome possibly indicating some predictive power, although generally the evidence was limited. The small number of studies that assessed each factor and heterogeneity among studies precluded a quantitative analysis of the results. Additional studies are necessary to further substantiate these associations before strong conclusion can be drawn.

INTRODUCTION

Lumbar radiculopathy is generally caused by lumbar disc herniation or spinal stenosis secondary to degenerative changes located in the lateral recess or in the neuroforamen [1]. Due to an intricate interplay of physical compression, inflammatory and immunological processes, irritation of a nerve root arises which may result in symptoms of leg pain, sensory and motor deficits [2, 3]. Since these symptoms usually interfere with daily activities, lumbar radiculopathy is often considered strongly invalidating and patients seek for effective treatments to relieve symptoms [4]. For lumbar disc herniation current guidelines recommend starting conservative therapy during the first weeks to months before switching to more invasive options as this condition is generally self-limiting over time [5-7]. For spinal stenosis there is no full consensus on treatment strategies as some patients may experience spontaneous relief of symptoms despite the general idea that spinal stenosis has a more chronic character than disc herniation. As a result, clinical guidelines are more incongruous with each other [8-11]. Nevertheless, for both conditions there are treatment options aside from surgery for patients whose symptoms do not resolve naturally within reasonable time.

Transforaminal epidural steroid injections (TEI) is a treatment that has been offered to patients with (unilateral) lumbar radiculopathy due to disc herniation or stenosis for decades [12]. These steroid injections are aimed at reducing inflammatory cytokines and downregulating immunological reactions in order to relieve pain symptoms [13, 14]. They are assumed to have a temporary analgesic effect but may result in complete resolution of symptoms in some patients. In addition to the transforaminal route, epidural steroid injections can be administered through a caudal or interlaminar approach, but the transforaminal injection is considered to be the most selective one. Despite the wide use of TEI in patients with lumbar radiculopathy due to disc herniation or spinal stenosis its effectiveness remains a matter of debate. Studies have demonstrated contradictory results with large variation in efficacy and success percentages, but it remains unclear which determinants are of influence on the efficacy of TEI [10, 12, 15-18]. Some studies have suggested that TEI effectiveness may be affected by the etiology of lumbar radiculopathy (i.e., disc herniation or stenosis), whereas other studies propose alternative subgroups of patients that may benefit more from treatment with TEI. The varying effectiveness of TEI across studies and the uncertainty of which patient groups benefit more than others have led to discrepancies between clinical guidelines on the applicability and timing of TEI [5, 6, 8, 12, 15, 18-20]. Therefore, establishing predictive factors for treatment success with TEI could aid in tailoring treatment strategies for

patients with lumbar radiculopathy, possibly obviating the need for surgery or expediting the timing of surgical intervention. For this reason, we aim to provide a comprehensive overview of the current literature on demographic, clinical and radiological factors and their association with TEI efficacy in patients with unilateral lumbar radiculopathy due to disc herniation or stenosis.

METHODS

This systematic review was conducted in accordance with the PRISMA guidelines.

Search and selection

Six databases (PubMed, Embase, Web of Science, Cochrane Library, Emcare and Academic Search Premier) were searched from inception to 25 July 2025. A search strategy was constructed and adapted to every database by an expert librarian (Appendix 1). Full texts as well as meeting abstracts were included in the search. Strings were included for prospective and retrospective studies on patients with lumbar disc herniation or degenerative stenosis with unilateral symptoms. Acceptable outcome measures were the Numerical Rating Scale (NRS) or the Visual Analogue Scale (VAS) for pain scores and the Oswestry Disability Index (ODI) or the Roland Morris Disability Questionnaire (RMDQ) for physical disability scores. All results were screened by two reviewers (EV and NV) separately based on title and, consequently, on abstract. The remaining full texts were evaluated by two reviewers independently. Any discrepancies were resolved by discussion or consulting a third reviewer (CVL).

Inclusion and exclusion criteria

Studies analyzing the association between predictive factors and the success of TEI were included if they met the inclusion and exclusion criteria. Studies were eligible if the study design was a Randomized Controlled Trial (RCT), prospective cohort or retrospective cohort and patients had unilateral radicular pain. The choice to only include patients with unilateral symptoms was based on the idea that patients with bilateral complaints may not be sufficiently helped with a single transforaminal injection and pain and disability scores would not represent the treatment efficacy accurately. In addition, studies were eligible if patients had MRI confirmed LDH or degenerative stenosis (central, lateral recess and/or foraminal), they were injected on only one level on one side, received TEI containing any type of steroids, the outcome measure was based on symptom relief, leg pain, functionality scores or need for subsequent surgery, the cohort consisted of at least 20 patients, the follow-up was at least two weeks, and the

article was written in English. Studies were excluded if they presented a case report or series, systematic review or meta-analysis, included patients with previous back surgery, scoliosis or spondylolisthesis, allowed for administration of multiple injections, used an injection approach other than the transforaminal route, evaluated technical aspects of transforaminal injections, evaluated different types or concentrations of corticosteroids, used non-fluoroscopy guided injections, only measured back pain, or investigated a predictive factor derived from invasive examinations that are not part of routine work-up (e.g., not every patient with lumbar radiculopathy undergoes EMG).

Risk-of-bias assessment

Risk of bias assessment was performed by two reviewers separately (EV and NV), using an adjusted version of the Cowley risk of bias criteria (Appendix 2) [21]. Each study could be awarded a minimum of 0 and maximum of 10 points. Studies with a score above seven points were classified as low risk of bias, studies with 5-7 points as intermediate risk of bias and studies with fewer than 5 points as high risk of bias. Differences between the two assessors were resolved during a consensus meeting with a third reviewer (CVL).

Data extraction and analysis

From all included studies, data on methods (study design, cohort size), patients (inclusion and exclusion criteria, baseline characteristics, diagnostic characteristics), treatment, outcome variables and results were gathered by one reviewer with a second reviewer verifying the final data extraction. When patients were dichotomized based on treatment outcome the definition of treatment success was included. Results were based on data from text, tables or graphs and included raw data, whether an association was found and the conclusion the authors had drawn. For continuous outcome measures means and standard deviation (SD) or median value and range were retrieved, whereas for dichotomized outcomes absolute numbers, percentages, or odds ratios (OR) were gathered. A qualitative analysis was performed for all results.

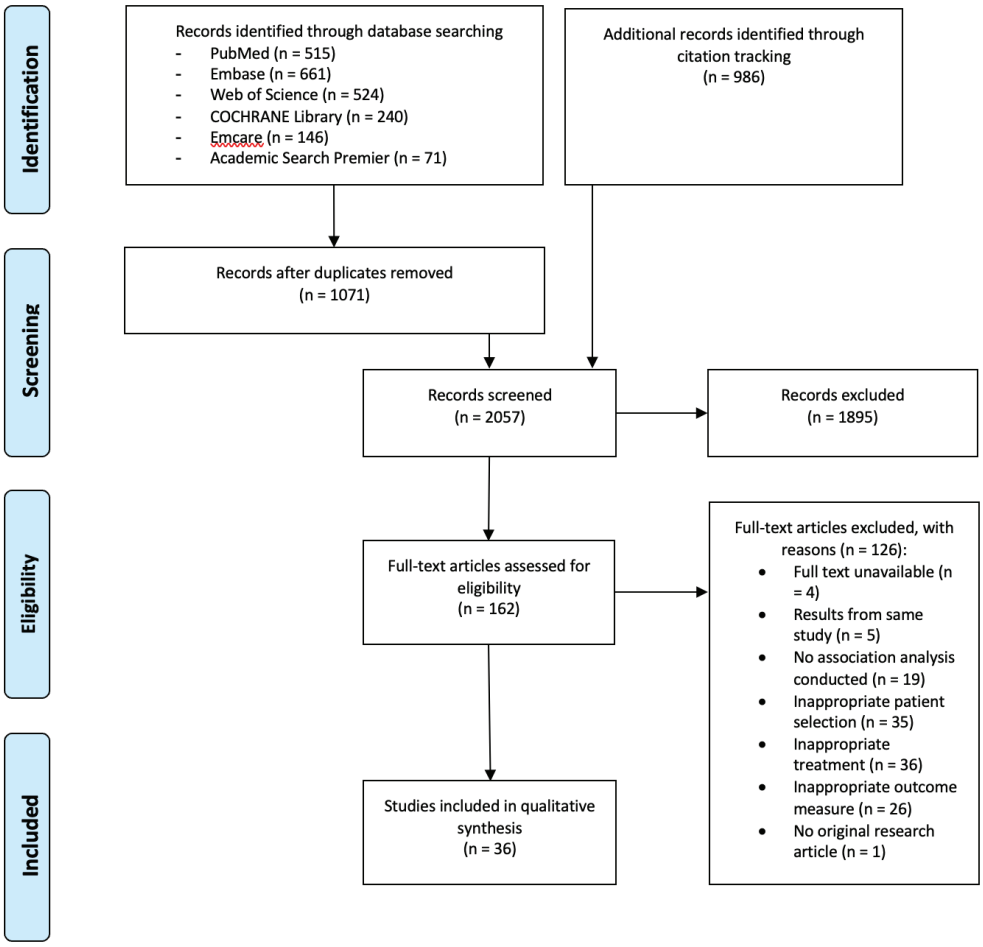
RESULTS

Article selection

The literature search yielded a total of 2057 unique articles of which 162 references were selected for full-text screening. Eventually, 36 articles were included for this review (Fig. 1) [22-57]. Twenty-two articles assessed patients with lumbar disc herniation, five articles evaluated patients with spinal stenosis and nine studies included both patient groups. Twenty-one studies were prospective,

and fifteen studies were conducted retrospectively. All studies were published between 2006 and 2025 and the patient population size ranged between 20 and 2024 patients. A more elaborate overview of study characteristics is given in Appendix 3.

Fig. 1 Flowchart of the article search and selection process. A total of 2057 records were screened of which 36 articles were included in this review.



Risk-of-bias assessment

Of the 36 reports, ten scored 8 or 9 out of ten points and were considered low risk of bias [28, 31, 32, 39, 45, 50, 52, 53, 55, 57]. A total of 22 reports were awarded 5 to 7 points and were categorized as intermediate risk of bias [22, 24-27, 29, 30, 35, 36, 38, 40-44, 46-49, 51, 54, 56]. Four studies scored four points or lower and were judged as high-risk of bias [23, 33, 34, 37] (Appendix 4).

Factors associated with patient outcome

In all studies, a total of 37 different factors were assessed for their association with patient outcome after treatment with a TEI. Of these 37 factors, eight variables were demographic, ten were clinical and nineteen were radiological. Results are summarized in Table 1. All extracted data are provided in Appendix 5.

Table 1 An overview of all demographic, clinical and radiological variables studied for their association with TEI outcome. For each variable the number of studies that found an association and the number of studies that did not find an association are provided.

Variable	Number of studies that found an association	References	Number of studies that did not find an association	References
Demographic factors				
Age	1	[37]	8	[22, 24, 29, 36, 38, 41, 46, 50]
BMI	1	[37]	4	[36, 40, 46, 50]
Duration of symptoms	4	[29, 35, 37, 48]	6	[24, 36, 38, 41, 46, 50]
Injection level	0		6	[22, 24, 42, 50, 56]
Injection side	0		4	[24, 37, 38, 56]
MRI to block	0		1	[38]
Sex	1	[37]	6	[24, 36, 38, 41, 42, 50]
Smoking status	0		1	[36]
Clinical factors				
hsCRP	0		1	[44]
Positive Slump test	2	[31, 32]	0	
Pre-injection functional score	2	[36, 37]	1	[45]
Pre-injection pain score	1	[37]	5	[29, 36, 45, 46, 50]
Presence of HACS	1	[47]	0	
Presence of neurologic deficit	1	[36]	1	[31]
Presence of neuropathic pain	1	[51]	0	
Vitamin D deficiency	1	[43]	0	
Post-injection functional score	1	[36]	1	[45]
Post-injection pain score	1	[50]	2	[36, 45]

Table 1 An overview of all demographic, clinical and radiological variables studied for their association with TEI outcome. For each variable the number of studies that found an association and the number of studies that did not find an association are provided. (*continued*)

Variable	Number of studies that found an association	References	Number of studies that did not find an association	References
Radiological factors				
Facet tropism	1	[27]	0	
Gadolinium enhancement of nerve root	1	[54]	0	
IVD signal intensity	1	[56]	0	
Lesion level	0		3	[24, 29, 36]
Lesion severity	2	[28, 39]	3	[24, 33, 36]
Lesion side	0		1	[36]
Location of LDH	2	[32, 34]	3	[29, 36, 56]
Grade of facet degeneration	1	[49]	0	
Grade of nerve root compression	3	[31, 32, 56]	4	[26, 40, 46, 50]
Mean spinal nerve intensity	1	[38]	0	
Morphology of LDH	2	[36, 55]	1	[30]
Nerve-to-fat signal ratio	1	[38]	0	
Paraspinal muscle area	0		1	[46]
Presence of Modic changes	0		1	[57]
Presence of transitional vertebrae	3	[36, 52, 53]	1	[50]
Radiological etiology of symptoms	3	[24, 25, 42]	3	[29, 31, 41]
Type of disc degeneration	1	[32]	2	[36, 56]
Type of Gadolinium enhancement of nerve root	0		1	[54]
Type of transitional vertebrae	1	[53]	0	

BMI: Body Mass Index; HACS: human assumed central sensitization; hsCRP: high-sensitivity C-reactive protein; IVD: intervertebral disc; LDH: lumbar disc herniation; MRI: Magnetic Resonance Imaging

Demographic factors

One high risk-of-bias study reported that age was positively associated with improvement in pain and disability scores following TEI [37], whereas eight others did not demonstrate a relationship [22, 24, 29, 36, 38, 41, 46, 50]. Duration of symptoms was negatively associated in four out of ten studies in terms of pain and functionality scores [29, 35, 37, 48], including one high-risk-of-bias study [37].

Sex was evaluated in seven studies [24, 36-38, 41, 42, 50], but only one (high risk of bias) observed better outcomes in men at two months follow-up [37]. Among fives studies investigating body mass index (BMI) [36, 37, 40, 46, 50], only a single high-risk-of-bias report found a marginal positive association with short-term functional improvement but not with pain or longer-term results [37].

For injection level [24, 38, 42, 50, 56], injection side [24, 37, 38, 56], the number of days between MRI and TEI treatment [38] and smoking status [36], no association was found in any of the studies.

Clinical factors

Six studies assessed pre-injection pain scores [29, 36, 37, 45, 46, 50] and three pre-injection functional scores [36, 37, 45]. Only one high-risk of bias study found an association between pre-injection pain scores and better short-term outcome [37]. That same study, along with another [36], reported an association between baseline functional scores and post-injection outcomes.

High-sensitivity C-reactive protein levels were not associated with TEI outcome [44]. One study showed that patients with neuropathic pain experienced greater functional improvement at three months than those without [51]. A positive Slump test correlated with greater reductions in pain and disability at three weeks [31, 32], whereas vitamin D deficiency and the presence of central sensitization were negatively associated with outcome [43, 47]. One study found no association between neurological deficit and treatment effect [31], while, contrastingly, another reported poorer outcomes in patients with sensory symptoms [36].

Pain reduction one hour post-injection predicted three-month success in one study [50], although two other studies found no such association [36, 45]. However, one of these did find an association between higher functionality at three weeks and one-year outcome [36].

Radiological factors

Three of six studies demonstrated an association between the radiological etiology (LDH or stenosis) and treatment effect [24, 25, 29, 31, 41, 42]. Patients with lumbar stenosis achieved greater immediate post-injection improvement than those with LDH in one study [24], whereas another reported superior long-term results for foraminal stenosis compared to LDH or spinal/lateral recess stenosis [42]. Contrastingly, the third study demonstrated better pain reduction in LDH than LSS at four weeks [25].

Type of disc degeneration (i.e., bulging, protrusion, extrusion) was assessed in three studies [32, 36, 56], but only one found greater short-term functional improvement for disc extrusion compared to disc protrusion [32]. Two of three studies evaluating disc morphology reported an association with TEI outcome, though one was high risk-of-bias [23, 55].

Findings on LDH location were inconsistent: three studies found no correlation [29, 36, 56], whereas two other studies suggested better outcomes for central/subarticular or paramedian herniations compared with foraminal/extraforaminal lesions [32, 34], one of which had high risk of bias.

Three studies investigated MRI signal characteristics. Increased nerve-to-fat signal ratio and mean spinal nerve intensity on axial T2-weighted images predicted better post-injection outcomes in one study [38], and gadolinium enhancement of the nerve root was also associated with improvement in another, independent whether the post-dorsal root ganglion enhanced [54]. High signal intensity zones within the intervertebral disc were associated with lower pain scores at two weeks, but not for success rates or for three-month post-injection outcomes [56].

Lesion level and side were not associated with TEI outcome [24, 29, 36]. Lesion severity (grading of LDH or stenosis) was examined in five studies using various grading scales [24, 28, 33, 36, 39]. Four categorized lesion severity as mild, moderate or severe: two reported better outcomes for mild-moderate or moderate stenosis compared to severe cases [28, 39], whereas two studies found no association for either LDH or stenosis patients [24, 33], though the latter study had a high risk of bias. Another study also found no correlation between Pfirrmann grade and post-injection outcome [36].

Seven studies evaluated the degree of nerve root compression using different grading systems. Three demonstrated outcome differences by compression grade. In two studies, high-grade subarticular compression was associated with

greater short-term functional improvement compared to high-grade foraminal or low-grade (subarticular) compression, whereas pain scores and outcomes between low- and high-grade foraminal compression did not differ [31, 32]. A third study reported lower pain scores at three months in patients with any compression versus none, but not at two weeks or in success-rate analyses [56]. By contrast, three studies using the (modified) Pfirrmann scale and one dichotomizing discoradicular contact found no association between compression grade and TEI outcome [26, 40, 46, 50].

Four studies assessed the presence of transitional vertebrae of which three found a negative association with TEI outcome [36, 52, 53]. Specifically, sacralization was associated with worse results than lumbarization or absence of transitional vertebrae [53]. Facet tropism and high-grade facet degeneration were associated with less favorable TEI responses [27, 49]. No associations were demonstrated for paraspinal muscle area or the presence of Modic changes [46, 57].

DISCUSSION

This systematic review summarizes the current literature on demographic, clinical and radiological variables, and their association with patient outcome after TEI. For demographic variables, there is some evidence for an association between duration of symptoms and outcome after TEI, with four out of ten studies indicating worse patient outcome when symptom duration was longer. Chronicity of inflammatory processes can result in pain perception and conduction alterations, and, therefore, may render local treatment with corticosteroids less effective [29, 37, 58, 59]. However, of these studies, three had an intermediate and one had a high risk of bias. Of the studies that found no association, five were intermediate and one was low risk of bias. As a result, there is no convincing evidence that can fully sustain this association and, thus, no implication that patients with chronic radicular symptoms should be advised against treatment with TEI.

In terms of clinical factors, a positive Slump test was positively associated with TEI outcome. A possible explanation for the positive association may be that this test makes a distinction between lumbar radiculopathy and leg pain symptoms without neural compromise, while TEI is only aiming to relieve radicular pain. However, the evidence was limited with only two studies reporting on this variable. In addition, five out of six studies found no association with pre-injection

pain score, indicating that both patients with mild and severe radicular pain may benefit from TEI.

Similarly, there was limited evidence for the association between stenosis rather than LDH patients and improved outcome after TEI. This appears to contrast with the results related to symptom duration in some studies, since, in general, patients with lumbar stenosis tend to suffer from chronic symptoms more frequently than LDH patients. However, it is possible that the inflammatory component plays a more significant role in stenosis patients compared to LDH patients. Some evidence was available for the association between the presence of transitional vertebrae and inferior TEI results. This association may be due to a more challenging identification of the affected nerve root to aim for with TEI, or to the technically more complicated procedure in applying TEI. Furthermore, there was limited evidence suggesting that more centrally located LDH are associated with better outcomes following TEI compared to more lateral locations. Central or paramedian herniated discs may be larger, potentially eliciting a higher release of inflammatory cytokines. Although this would also result in more compression, lateral herniations may actually exert greater physical impingement due to the restricted space available for the nerve root within the foramen. Hence, in the case of central and paramedian disc herniations, the inflammatory component rather than mechanical compression may affect the nerve root more, leading to a higher response to TEI. Conversely, three out of seven studies demonstrated a positive correlation between higher grades of nerve root compression and superior TEI outcomes. This finding appears counterintuitive, as patients with less physical compression are generally thought to have a larger inflammatory or immunological component and would therefore be expected to respond better to TEI. However, the reliability of this association is limited: it was observed in fewer than half of the available studies and only for functional improvement, not for pain scores [31, 32]. Moreover, one study included only patients with either marked (≥ 10-point) improvement in ODI or no response, thereby excluding those with intermediate outcomes [31], while another found a statistically but not clinically significant difference [56].

A strength of the current review is that strict inclusion and exclusion criteria were applied during the literature search. Only studies that described treatment with a single transforaminal epidural steroid injection were eligible for inclusion in order to maximize the comparability between treatments. Moreover, only studies on patients with unilateral radiculopathy could be included as it was hypothesized that TEI may not be as effective in patients with bilateral complaints due to its route of administration, hence our results would be confounded by this. Furthermore, reports using outcome instruments for pain symptoms other

than NRS and/or VAS, or for functional disability other than ODI and/or RMDQ scores were excluded for comparability purposes. However, our strict inclusion and exclusion criteria may have resulted in the omission of articles that held interesting results but not fully met our criteria. Another limitation is that comparability of studies may be restrained by the diversity of steroids, volumes and dosages used for TEI, and the variety in follow-up moments across studies. This may have affected the results from our review.

To our best knowledge, only three other reviews have been published before that evaluated the association of factors with outcome after epidural steroid injections [60-62], though they had different approaches than we chose for the current paper. One review focused on RCTs of epidural steroid injection versus placebo, different epidural injection approaches, different type and dose of corticosteroids, and yielded only limited evidence for epidural corticosteroid efficacy, without appointing clear prognostic factors [60]. Another review considered laboratory markers and imaging characteristics as predictive tools for epidural steroid injection efficacy [61], and concluded that there is some evidence for an association of more pain relief in patients with more nerve root compression. Additionally, they reported that IFN-γ obtained from epidural lavage fluid predicts short-term pain reduction, although it was mentioned that the use of this assay is challenging and expensive in a clinical setting. Yet, for that particular review studies on all epidural techniques and both patients with cervical and lumbar radicular pain were included. A more recent review on prognostic factors in disc-related sciatica reported findings consistent with ours, demonstrating some evidence of a negative association between symptom duration and TEI outcome, as well as for the degree of nerve root compression, for which our review provided only limited support [62]. However, the authors accepted studies on patients with bilateral sciatica, all three epidural techniques with or without image-guidance, multi-level injections, they did not include the PubMed database, excluded stenosis patients and included fewer prospective studies.

The heterogeneity with regard to design and methodology between studies complicates the comparability of findings and may explain that some results are contradictory. In order to identify subgroups of LDH and stenosis patients that benefit more from TEI treatment, future studies should be specifically designed for development of a prediction model for TEI success. Moreover, future studies should focus on predictive factors for the effect of a single injection rather than a series of injections.

Overall, there is no strong evidence for any association between the assessed variables and TEI outcome based on the studies in this review. There is limited support that duration of symptoms, positive Slump test, radiological etiology of symptoms, location of LDH, transitional vertebrae and nerve root compression may influence TEI efficacy, but these associations need to be further substantiated in future studies before firm conclusions can be drawn.

Appendices will be available online after publication.

REFERENCES

1. Tarulli AW, Raynor EM. Lumbosacral radiculopathy. Neurol Clin. 2007;25(2):387-405.
2. Kobayashi S, Yoshizawa H, Yamada S. Pathology of lumbar nerve root compression. Part 1: Intraradicular inflammatory changes induced by mechanical compression. J Orthop Res. 2004;22(1):170-9.
3. Robinson JP, Brown PB, Fisk JD. Pathophysiology Of Lumbar Radiculopathies and The Pharmacology Of Epidural Corticosteroids and Local Anesthetics. Physical Medicine and Rehabilitation Clinics of North America. 1995;6(4):671-90.
4. Winter CC, Brandes M, Müller C, Schubert T, Ringling M, Hillmann A, et al. Walking ability during daily life in patients with osteoarthritis of the knee or the hip and lumbar spinal stenosis: a cross sectional study. BMC Musculoskelet Disord. 2010;11:233.
5. Kreiner DS, Hwang SW, Easa JE, Resnick DK, Baisden JL, Bess S, et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. Spine J. 2014;14(1):180-91.
6. Lorio M, Kim C, Araghi A, Inzana J, Yue JJ. International Society for the Advancement of Spine Surgery Policy 2019-Surgical Treatment of Lumbar Disc Herniation with Radiculopathy. Int J Spine Surg. 2020;14(1):1-17.
7. Chiu CC, Chuang TY, Chang KH, Wu CH, Lin PW, Hsu WY. The probability of spontaneous regression of lumbar herniated disc: a systematic review. Clin Rehabil. 2015;29(2):184-95.
8. Kreiner DS, Shaffer WO, Baisden JL, Gilbert TJ, Summers JT, Toton JF, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). Spine J. 2013;13(7):734-43.
9. Kreiner S, Shaffer W, Baisden J, et al. *Evidence-Based Clinical Guidelines for Multidisciplinary Spine Care. Diagnosis and Treatment of Degenerative Lumbar Spinal Stenosis*. North American Spine Society; 2011.
10. Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and Management of Lumbar Spinal Stenosis: A Review. Jama. 2022;327(17):1688-99.
11. Anderson DB, Luca K, Jensen RK, Eyles JP, Van Gelder JM, Friedly JL, et al. A critical appraisal of clinical practice guidelines for the treatment of lumbar spinal stenosis. Spine J. 2021;21(3):455-64.
12. Palmer WE. Spinal Injections for Pain Management. Radiology. 2016;281(3):669-88.
13. McLain RF, Kapural L, Mekhail NA. Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J. 2005;5(2):191-201.
14. Starkweather A, Witek-Janusek L, Mathews HL. Neural-immune interactions: implications for pain management in patients with low-back pain and sciatica. Biol Res Nurs. 2005;6(3):196-206.
15. Deer TR, Grider JS, Pope JE, Lamer TJ, Wahezi SE, Hagedorn JM, et al. Best Practices for Minimally Invasive Lumbar Spinal Stenosis Treatment 2.0 (MIST): Consensus Guidance from the American Society of Pain and Neuroscience (ASPN). J Pain Res. 2022;15:1325-54.
16. Helm Ii S, Harmon PC, Noe C, Calodney AK, Abd-Elsayed A, Knezevic NN, et al. Transforaminal Epidural Steroid Injections: A Systematic Review and Meta-Analysis of Efficacy and Safety. Pain Physician. 2021;24(S1):S209-s32.

17. Smith CC, McCormick ZL, Mattie R, MacVicar J, Duszynski B, Stojanovic MP. The Effectiveness of Lumbar Transforaminal Injection of Steroid for the Treatment of Radicular Pain: A Comprehensive Review of the Published Data. Pain Med. 2020;21(3):472-87.
18. Services Department of Health and Human Services. *Local Coverage Determination* - Epidural Steroid Injections for Pain Management. 2023. Available from: https://www.cms.gov/medicare-coverage-database/view/lcd.aspx?LCDId=39054. [Accessed on: July 12, 2023].
19. Comer C, Ammendolia C, Battié MC, Bussières A, Fairbank J, Haig A, et al. Consensus on a standardised treatment pathway algorithm for lumbar spinal stenosis: an international Delphi study. BMC Musculoskelet Disord. 2022;23(1):550.
20. Manchikanti L, Knezevic NN, Navani A, Christo PJ, Limerick G, Calodney AK, et al. Epidural Interventions in the Management of Chronic Spinal Pain: American Society of Interventional Pain Physicians (ASIPP) Comprehensive Evidence-Based Guidelines. Pain Physician. 2021;24(S1):S27-s208.
21. Cowley DE. Prostheses for primary total hip replacement. A critical appraisal of the literature. Int J Technol Assess Health Care. 1995;11(4):770-8.
22. Abdullah A, Haque SSM, Pathan AA. Pain relief outcomes after selective nerve root block in patients with lumbar disc herniation. International Journal of Life Sciences Biotechnology and Pharma Research. 2025;14(4):333-7.
23. Arun-Kumar K, Jayaprasad S, Senthil K, Lohith H, Jayaprakash KV. The Outcomes of Selective Nerve Root Block for Disc Induced Lumbar Radiculopathy. Malays Orthop J. 2015;9(3):17-22.
24. Bang YJ, Park HJ, Sim WS, Lee DW, Lee JY. Correlation between paravertebral spread of injectate and clinical efficacy in lumbar transforaminal block. Scientific Reports. 2020;10(1):11508.
25. Bokhari SHR, Awais MZ, Iqbal MT. Four-Week Outcomes of Transforaminal Epidural Injections in Patients With Lumbar Radicular Pain. Cureus. 2024;16(12):e75929.
26. Budrovac D, Radoš I, Hnatešen D, Haršanji-Drenjančević I, Tot OK, Katić F, et al. Effectiveness of Epidural Steroid Injection Depending on Discoradicular Contact: A Prospective Randomized Trial. Int J Environ Res Public Health. 2023;20(4).
27. Celenlıoglu AE, Sencan S, Gunduz OH. Does facet tropism negatively affect the response to transforaminal epidural steroid injections? A prospective clinical study. Skeletal Radiol. 2019;48(7):1051-8.
28. Chang MC, Lee DG. Outcome of Transforaminal Epidural Steroid Injection According to the Severity of Lumbar Foraminal Spinal Stenosis. Pain Physician. 2018;21(1):67-72.
29. Cyteval C, Fescquet N, Thomas E, Decoux E, Blotman F, Taourel P. Predictive factors of efficacy of periradicular corticosteroid injections for lumbar radiculopathy. AJNR Am J Neuroradiol. 2006;27(5):978-82.
30. Dhandapani K, Som D, Muthiahpandian P, Miller A, Venkatesan A, Baid M, et al. Functional Outcomes and Successful Predictors of Lumbar Transforaminal Epidural Steroid Injections (LTFESIs) for Lumbar Radiculopathy Under Fluoroscopic Guidance: A Prospective Study. Cureus. 2023;15(12):e50257.
31. Ekedahl H, Jönsson B, Annertz M, Frobell RB. Three week results of transforaminal epidural steroid injection in patients with chronic unilateral low back related leg pain: The relation to MRI findings and clinical features. J Back Musculoskelet Rehabil. 2016;29(4):693-702.

32. Ekedahl H, Jönsson B, Annertz M, Frobell RB. The 1-Year Results of Lumbar Transforaminal Epidural Steroid Injection in Patients with Chronic Unilateral Radicular Pain: The Relation to MRI Findings and Clinical Features. Am J Phys Med Rehabil. 2017;96(9):654-62.
33. Elsayad E, Mortada M, Elhawala TAES, Hady EAE, Hijab A. SAT0596 Efficacy of epidural steroid injection in lumbar spinal stenosis is not related to the degree of severity by mri. Annals of the Rheumatic Diseases. 2017;76(Suppl 2):1000-.
34. Guclu B, Deniz L, Yuce Y, Adilay U, Aytar H, Turkoglu M, et al. Transforaminal Epidural Steroid Injection in the Treatment of Pain in Foraminal and Paramedian Lumbar Disc Herniations. Turk Neurosurg. 2020;30(3):394-9.
35. Guclu B, Deniz L, Katar S, Ozdemir A. Optimal Timing and Outcome of Transforaminal Epidural Steroid Injection for the Management of Radicular Pain due to Extruded Lumbar Disc Herniation. Turk Neurosurg. 2023;33(3):509-13.
36. Kanna RM, Shetty AP, Rajasekaran S. Predictors of Successful Outcomes of Selective Nerve Root Blocks for Acute Lumbar Disc Herniation. Global Spine J. 2019;9(5):473-9.
37. Kaufmann TJ, Geske JR, Murthy NS, Thielen KR, Morris JM, Wald JT, et al. Clinical effectiveness of single lumbar transforaminal epidural steroid injections. Pain Med. 2013;14(8):1126-33.
38. Kim YH, Kim HJ, Seo J, Chai JW, Song HG, Choi YH, et al. Spinal nerve signal intensity on Dixon T2-weighted water-only sequence: an important outcome predictor after lumbar transforaminal epidural injection. Eur Radiol. 2021;31(12):9459-67.
39. Kim M, Bak J, Goh D, Bae J, Shin K, Son HJ, et al. Changes in pain scores and walking distance after transforaminal epidural steroid injection in patients with lumbar foraminal spinal stenosis. Medicine (United States). 2023;102(25):E34032.
40. Kokar S, Sencan S, Yildiz B, Polat BA, Sacaklidir R, Gunduz OH. The Effect of Obesity on the Treatment Outcomes of Lumbar Transforaminal Epidural Steroid Injections. Turk Neurosurg. 2025;35(1):120-8.
41. Lee JW, Kim SH, Lee IS, Choi JA, Choi JY, Hong SH, et al. Therapeutic effect and outcome predictors of sciatica treated using transforaminal epidural steroid injection. AJR Am J Roentgenol. 2006;187(6):1427-31.
42. Olguner SK, Celiktas M, Oktay K, Arslan A, Bilgin E, Gezercan Y, et al. Comparison of 1-year results of single transforaminal epidural steroid injection among patients with different spinal pathologies-related radicular pain. Niger J Clin Pract. 2020;23(6):835-41.
43. Ozturk EC, Sencan S, Sacaklidir R, Albayrak O, Gunduz OH. The Impact of Vitamin D Deficiency to Treatment Success of Transforaminal Epidural Steroid Injection. Pain Physician. 2021;24(5):E619-e24.
44. Park CH, Lee SH. Prognostic usefulness of high sensitivity C-reactive protein for transforaminal epidural steroid injection in patients with radicular pain. Pain Med. 2011;12(2):219-23.
45. Ploumis A, Christodoulou P, Wood KB, Varvarousis D, Sarni JL, Beris A. Caudal vs transforaminal epidural steroid injections as short-term (6 months) pain relief in lumbar spinal stenosis patients with sciatica. Pain Med. 2014;15(3):379-85.
46. Saçaklıdır R, Soydemir E, Şencan S, Gündüz OH. The effect of lumbar multifidus cross-sectional areas on transforaminal epidural steroid injection: An observational clinical study. Agri. 2023;35(4):205-11.

47. Sahin T, Sacaklidir R, Sancar M, Öztürk EC. The effect of human assumed central sensitization on transforaminal epidural steroid injection in chronic lumbar radiculopathy: An observational study. J Back Musculoskelet Rehabil. 2024;37(6):1749-55.
48. Sariyildiz MA, Batmaz İ, Yazmalar L, Güneş M, Turan Y. The effectiveness of transforaminal epidural steroid injections on radicular pain, functionality, psychological status and sleep quality in patients with lumbar disc herniation. J Back Musculoskelet Rehabil. 2017;30(2):265-70.
49. Sekizkardes Tutuncu M, Sencan S, Kurt CB, Kokar S, Gunduz OH. The role of facet joint degeneration in the treatment success of transforaminal epidural steroid injection: a retrospective clinical study. Skeletal Radiol. 2025;54(8):1689-96.
50. Şencan S, Çelenlioğlu AE, Asadov R, Gündüz OH. Predictive factors for treatment success of transforaminal epidural steroid injection in lumbar disc herniation-induced sciatica. Turk J Med Sci. 2020;50(1):126-31.
51. Sencan S, Celenlioglu AE, Yazici G, Gunduz OH. Transforaminal Epidural Steroid Injection Improves Neuropathic Pain in Lumbar Radiculopathy: A Prospective, Clinical Study. Neurol India. 2021;69(4):910-5.
52. Sencan S, Azizov S, Celenlioglu AE, Bilim S, Gunduz OH. Effect of sacralization on the success of lumbar transforaminal epidural steroid injection treatment: prospective clinical trial. Skeletal Radiol. 2022.
53. Son KM, Lee SM, Lee GW, Ahn MH, Son JH. The Impact of Lumbosacral Transitional Vertebrae on Therapeutic Outcomes of Transforaminal Epidural Injection in Patients with Lumbar Disc Herniation. Pain Pract. 2016;16(6):688-95.
54. Tak HJ, Jones R, Cho YW, Kim EH, Ahn SH. Clinical evaluation of transforaminal epidural steroid injection in patients with gadolinium enhancing spinal nerves associated with disc herniation. Pain Physician. 2015;18(2):E177-85.
55. Tanyel Saraçoğlu T, Erken B. Impact of Transforaminal Epidural Steroid Injection on Pain and Disability Outcomes by Lumbar Intervertebral Disc Herniation Class: A Prospective Study. Pain Med. 2025.
56. Tecer D, Adiguzel E, Tan AK, Taskaynatan MA. Role of Magnetic Resonance Imaging in Ascertaining the Success of Transforaminal Epidural Steroid Injection for Lumbar Radicular Pain. Pain Med. 2017;18(4):645-50.
57. Verheijen EJA, van der Vlist NRE, Bartels EC, van Haagen O, Vleggeert-Lankamp CLA. The effect of a transforaminal epidural injection in patients with lumbar disc herniation is not correlated with the presence of type II modic changes. Brain Spine. 2025;5:104222.
58. Rados I, Sakic K, Fingler M, Kapural L. Efficacy of interlaminar vs transforaminal epidural steroid injection for the treatment of chronic unilateral radicular pain: prospective, randomized study. Pain Med. 2011;12(9):1316-21.
59. Latremoliere A, Woolf CJ. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. J Pain. 2009;10(9):895-926.
60. Chou R, Hashimoto R, Friedly J, Fu R, Bougatsos C, Dana T, et al. Epidural Corticosteroid Injections for Radiculopathy and Spinal Stenosis: A Systematic Review and Meta-analysis. Ann Intern Med. 2015;163(5):373-81.
61. Benny BV, Patel MY. Predicting epidural steroid injections with laboratory markers and imaging techniques. Spine J. 2014;14(10):2500-8.
62. Nagington A, Foster NE, Snell K, Konstantinou K, Stynes S. Prognostic factors associated with outcome following an epidural steroid injection for disc-related sci-

atica: a systematic review and narrative synthesis. Eur Spine J. 2023;32(3):1029-53.

5

Transforaminal epidural injection versus continued conservative care in acute sciatica (TEIAS trial): study protocol for a randomized controlled trial

E.J.A. Verheijen, A.G. Munts, O.B.H.A.M. van Haagen, D. de Vries, O. Dekkers, W. van den Hout, C.L.A. Vleggeert-Lankamp

BMC Neurol 2019

ABSTRACT

Background

Sciatica is a condition that is characterised by radicular pain in the leg and primarily caused by a herniated lumbar intervertebral disk. In addition to leg pain, patients can experience back pain, leg numbness and leg weakness resulting in decreased productivity and social activity. The majority of sciatica cases recovers spontaneously and therefore patients are initially treated conservatively with oral pain medication. However, some patients experience intractable pain that severely impedes them and no consensus exists on the optimal conservative treatment to reduce this discomfort in the acute phase of sciatica. The aim of the TEIAS trial is to assess the effectiveness, cost-effectiveness and predictive capability on patient outcome of transforaminal epidural injection (TEI) compared to treatment with standard pain medication.

Methods

This study is designed as a prospective, open-label, mono-centered, randomized controlled trial. Patients that visit their general practitioner with complaints of radicular leg pain and meet the selection criteria are asked to participate in this study. Eligible patients will be randomized to treatment with TEI or to treatment with standard oral pain medication. Treatment of TEI will comprise lidocaine with methylprednisolone acetate for L3 and below and lidocaine with dexamethasone above L3. A total of 142 patients will be recruited and follow-up will occur after 1, 2, 4, 10 and 21weeks for assessment of pain, functionality, patient received recovery and cost-effectiveness. The primary outcome will be the average score for leg pain at 2 weeks. For this outcome we defined a clinically relevant difference as 1.5 on the 11-point NRS scale.

Discussion

Adequate conservative treatment in the acute phase of sciatica is lacking, particularly for patients with severe symptoms. Focusing on effectiveness, cost-effectiveness and predictive capability on patient outcome of TEI will produce useful information allowing for more lucid decision making in the conservative treatment of sciatica in the acute phase.

BACKGROUND

Sciatica is a condition that is characterised by radicular pain in the leg and generally caused by a herniated lumbar intervertebral disk. The herniated disk exerts pressure onto a lumbar nerve root that follows its way into the sciatic nerve. In addition to leg pain, this can lead to back pain, muscle weakness or muscle numbness resulting in decreased productivity and social activity [1–5]. This condition is regularly reported in GP practices and a lifetime prevalence around 40% has been recorded [6]. Together with other back pain-related conditions sciatica has a vast impact on the economy. In 2007, total costs of patients with lower back pain were estimated to be 3.5 billion euros in the Netherlands of which 88% was contributable to absenteeism and loss of productivity [7].

The prognosis of sciatica is favourable since the majority of cases recovers spontaneously. Therefore, patients are initially treated conservatively with standard oral pain medication unless cauda equine syndrome or progressive paresis is reported. Those groups of patients are directly referred to the neurologist or neurosurgeon. An earlier RCT demonstrated that surgical therapy resulted in a similar outcome compared to conservative therapy at 26 weeks follow-up and 40% of the patients assigned to the conservative group crossed over to the surgical group with a mean delay of 15 weeks [8]. Due to this knowledge it is now common to only offer surgery after 8 weeks of conservative care and preferably extend this time period to a total of 14–16 weeks in a process of Shared Decision Making [9, 10]. After this period of 'wait-and-see' patients can opt for surgery if symptoms have not ameliorated sufficiently. However, during the first weeks it can be challenging to sufficiently control the patient's pain and adequate treatment to reduce this discomfort in the acute phase of sciatica is yet to be found.

Transforaminal epidural injections (TEI) therapy might be the answer as it is postulated to have an analgesic and anti-inflammatory effect which can last several days to weeks. Analgesics and/or corticosteroids are deposited through a transforaminal approach in close proximity of the affected nerve root. Theoretically, this treatment would enable the patient to remain physically active while awaiting spontaneous recovery. Although its use is widespread, some patients experience a beneficial effect while others report no effect at all or only initial pain relief followed by recurrence of pain after several hours, days or weeks [11, 12]. In the Netherlands some hospitals offer TEI as standard care for sciatica but usually only after 14–16 weeks of conservative therapy.

Due to limited literature on TEI in patients with acute sciatica, it is still unclear whether TEI is a more (cost-)efficacious type of conservative treatment than oral pain medication in the acute phase. It remains difficult to differentiate between responders and non-responders and suggestions that response to TEI can be a predictor for long-term patient outcome require further research [12–14]. This study has the potential to answer relevant questions about TEI as minimally invasive conservative treatment in acute sciatica, assessing effectiveness on the short term and long term, predictive capability and cost-effectiveness.

DESIGN AND METHODS

Study design

The TEIAS (Transforaminal Epidural Injection in Acute Sciatica) Trial is designed as a prospective, open-label, mono-centred, randomized controlled trial. Eligible patients will be randomized to treatment with TEI (intervention group) or to treatment with standard oral pain medication (control group). Patients will be followed for a total of 21 weeks.

Aims and objectives

The aim of the TEIAS Trial is to determine whether transforaminal epidural injections are a more efficacious treatment to alleviate symptoms in patients suffering from acute sciatica than the current standard treatment with oral pain medication.

The primary objective is to estimate:

1. The average score on leg pain at 2 weeks after treatment with TEI compared to usual conservative care.

Secondary objectives are to estimate:

2. The duration of effect from TEI;
3. The percentage of patients that have experienced a satisfactory decrease in leg pain and increase in functionality at 4 weeks after treatment with TEI compared to usual conservative care;
4. The response to TEI at 2 weeks as a predictor of leg pain at 14-16 weeks and 26 weeks after onset of symptoms;
5. The correlation between response in leg pain, functionality and patient satisfaction at 1 week and at 2 weeks after randomization and at 10 and 21 weeks after randomization;

6. The correlation between baseline data and responders and non-responders to TEI after 2 weeks;
7. Cost-effectiveness of TEI to alleviate symptoms for acute sciatica;

Study setting and patient recruitment

This study will take place at the Spaarne Gasthuis (SG), Hoofddorp and Haarlem, the Netherlands in cooperation with researchers from the Department of Neurosurgery, Leiden University Medical Center (LUMC), Leiden, the Netherlands. Patient recruitment will occur at general practitioner offices that are part of the GP collective Haarlemmermeer and Kennemerland which embodies a total of 185 GP's. Patient inclusion will continue until the target sample size has been reached.

Patient enrolment

Patients suffering from acute sciatica that visit their general practitioner will be referred for participation in this study if they have a Numerical Rating Score (NRS) for leg pain of 6 or more on a 11-point scale and symptom duration of 3–8 weeks. Patients are contacted by the study investigator and inclusion and exclusion criteria are checked to determine whether the patient is eligible. Furthermore, it will be ascertained that the patient has access to e-mail in order to fill in the digital questionnaires at home. If patients are willing to cooperate and submit their informed consent to the researcher, they are referred to the neurosurgeon. The neurosurgeon will examine the patient and if the neurosurgeon is convinced that the patient is suffering from sciatica, the patient is enrolled by the researcher and randomized to receive TEI from the anaesthesiologist or to continue usual conservative care (oral pain medication). A complete description of inclusion and exclusion criteria can be found in Table 1.

Table 1 Patient inclusion and exclusion criteria

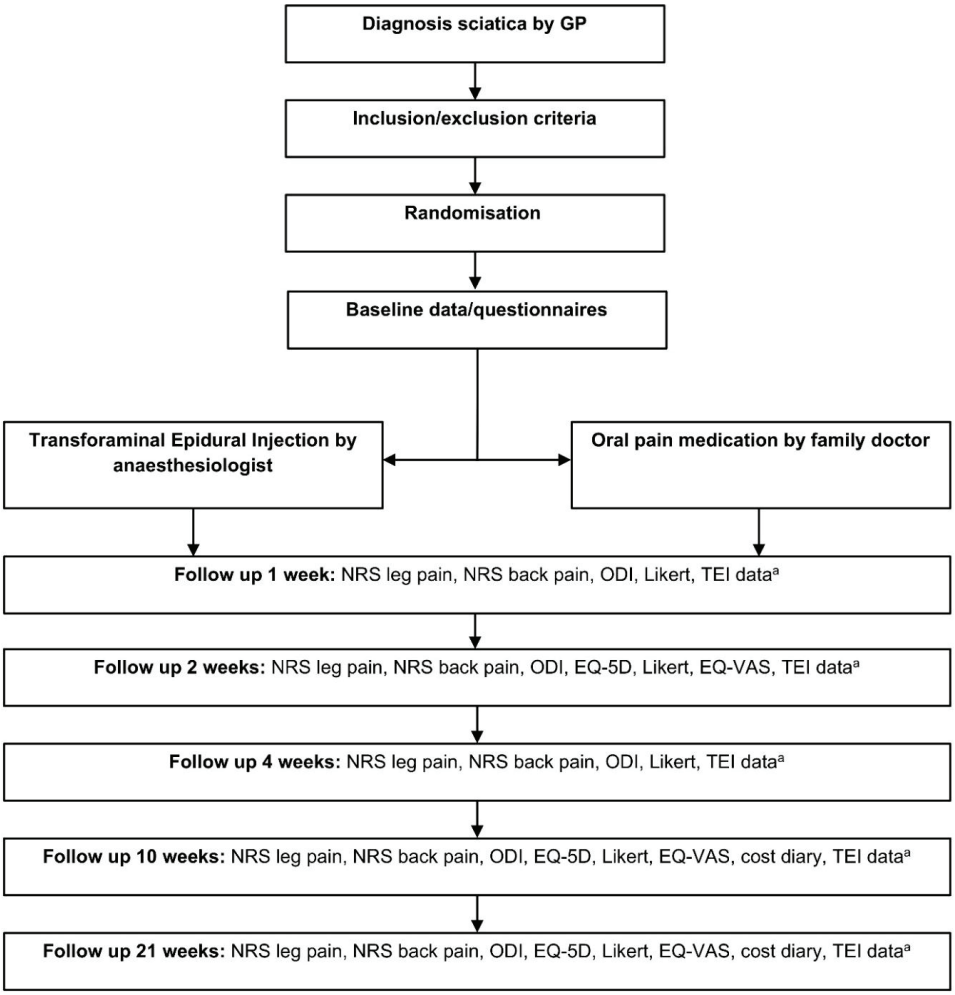
Inclusion criteria	Exclusion criteria
Diagnosed with sciatica by GP	Age under 18 years or above 55 years
NRS leg pain of 6 or more points on a 11-point scale with a duration of >3 and <8 weeks Confirmation of sciatica by neurosurgeon	Condition preventing to receive TEI Severe scoliosis
Signed informed consent	TEI received in 6 months before randomization date Surgery for sciatica at the same level
	Surgery for sciatica at another level within one year before inclusion
	Pregnancy

Description of processes

If the patient is eligible and decides to participate, block randomization in a 1:1 allocation ratio will determine which treatment the patient will follow. Randomization is done by using web-based randomization algorithm software (Castor EDC, Ciwit B.V., Amsterdam, The Netherlands). Details on blocking of randomization are unavailable to the researcher that will perform allocation. Consequently, baseline data will be collected, questionnaires will be filled in and the patient will receive instructions for follow-up questionnaires to be filled in at home. If the patient is randomized to the intervention group, the injection will be received within 4 working days after randomization. Level of injection will be determined by the anaesthesiologist although the neurosurgeon can advise. Patients from the control group will receive oral pain medication according to the care of their GP. Patients will be evaluated at baseline and after 1, 2, 4, 10 and 21 weeks. A schematic overview of the trial is shown in Fig. 1.

Patients that are randomized to the TEI group can receive additional injections if deemed profitable by the anaesthesiologist but an injection is only repeated after a certain time interval according to anaesthesiology guidelines for usual care. For all injections, details on timing, frequency and prevailing complications will be gathered.

Fig. 1 Flow diagram of TEIAS study procedures.



Description of interventions

Transforaminal epidural injection (intervention group)

Patients allocated to the intervention group will receive TEI within 4 working days after randomization. Before this procedure, the patient will undergo an X-ray of the lumbar spine in order to identify any spinal abnormalities that may complicate the treatment. Administration of TEI will be performed by an experienced anaesthesiologist. The patient will lie in a prone position on the table. The skin is sterilized with chlorhexidine. Under X-ray guidance using the "tunnel vision" technique the needle is placed via a transforaminal approach in close proximity of the nerve root and contrast agent is injected to confirm correct

positioning of the needle. In accordance with current Dutch anaesthesiologic guidelines, injections at L3 and below will contain 1,5 ml lidocaine 2% and 40 mg methylprednisolone acetate whereas injections above L3 will contain 1,5 ml lidocaine 1% and 10 mg dexamethasone. This is due to the possibility that particulate steroids like methylprednisolone acetate can occlude arterial blood vessels and result in an infarction of the spinal cord, the brain stem, cerebrum or cerebellum and therefore injections above L3 contain a non-particulate corticosteroid [15]. After the procedure, the patient will stay for 30 min at the recovery room for monitoring. If during the course of the study patients request additional injections, the complete procedure will be repeated for every injection.

Continued conservative care (control group)

The current standard for conservative care in patients with sciatica during the first 14 to 16 weeks is oral pain medication. Patients randomized to the control group can use over the counter medication or pain medication prescribed by the GP according to national guidelines. Use of pain medication will be recorded through the web-based questionnaires. If deemed necessary, the GP can additionally prescribe physiotherapy.

Referral to neurologist or neurosurgeon

If leg pain remains severe at 2 weeks after randomization, the patient is allowed to deviate from the study protocol. The GP can refer the patient to the neurologist and the patient can be offered surgery if applicable. An NRS for leg pain of 4 or higher makes the patient eligible for surgery if the patient requests surgery. Patients are still encouraged to postpone surgery till 14–16 weeks after onset of pain. This decision is made using Shared Decision Making and it is possible that the patient is operated before this time point. However, patients that are randomized cannot receive surgery within the first two weeks, unless cauda equine syndrome or progressive paresis is reported. It is expected that during these two weeks the transforaminal epidural injection will have its effect.

Table 2 Data collection schedule

	Baseline	**Weeks of follow-up**				
		1	**2**	**4**	**10**	**21**
Demographic data	✓					
NRS leg pain	✓	✓	✓	✓	✓	✓
NRS back pain	✓	✓	✓	✓	✓	✓
ODI	✓	✓	✓	✓	✓	✓
EQ-5D	✓		✓		✓	✓

Table 2 Data collection schedule (continued)

	Baseline			Weeks of follow-up		
Likert score		✓	✓	✓	✓	✓
EQ-VAS	✓		✓		✓	✓
Cost diary					✓	✓
TEI data[a]		✓	✓	✓	✓	✓

[a] Only collected for subjects that received TEI

Adjuvant care

Intake of oral pain medication is allowed in addition to assigned treatment. Since it is a randomized controlled trial, it is reasonable to expect that participants in both treatment arms will demonstrate comparable variation in oral pain medication.

Description of study parameters

Study parameters are assessed at baseline by a researcher and consequently by patients through online questionnaires at home. Data on the administration of TEI is collected by the anaesthesiologist. A summary of all study parameters and correlating follow-up moments can be found in Table 2.

Primary study parameter

- The primary study parameter, NRS leg pain, will be measured at all follow-up visits: at baseline and at 1, 2, 4, 10 and 21 weeks after randomization. NRS leg pain will be scored on a 0 to 10-point scale with increments of 1 point and patients are not entitled to see pain scores from previous follow-up moments. Each patient will score the NRS leg pain based on the mean leg pain over the past week.

Secondary study parameters

In order to assess functionality, perceived recovery, perceived general well-being and cost-effectiveness secondary study parameters will be measured.

- NRS back pain will be scored at all follow-up visits: at baseline and at 1, 2, 4, 10 and 21 weeks after randomization. NRS back pain will be measured on a 0 to 10-point scale with increments of 1 point and patients are not entitled to see pain scores from previous follow-up moments. Each patient will score the NRS back pain based on the mean back pain over the past week;
- The Oswestry Disability Index (ODI) will be measured to assess the functionality of the patient directed at walking and daily activities. The ODI will be scored on a 0 to 50-point scale at all follow-up visits: at baseline and at 1, 2,

4, 10 and 21 weeks after randomization. Patients are not entitled to see ODI scores from previous follow-up moments;
- Perceived recovery will be measured using the Likert scale. This is a 7-point scoring scale that ranges from 'completely recovered' to 'worse than ever'. The Likert score will be determined at baseline and at 1, 2, 4, 10 and 21 weeks follow-up;
- As an indicator for the patients' perspective on his or her health state, EuroQoL Visual Analogue Scale (EQ-VAS) will be determined using a 0 (as bad as death) to 100-point scale (perfect health) with increments of 1 point. This will be measured at baseline and at 2, 10 and 21 weeks follow-up;
- Data on timing of the injection (days after randomization), frequency of the injection, NRS leg pain after injection, prevailing complications and certainty of the anaesthesiologist that the patient is suffering from sciatica and that the injection has been administered at the correct lumbar level will be gathered during the total follow-up period of 21 weeks;
- For the economic evaluation (CEA), the five-level EuroQoL (EQ-5D) will be used. This questionnaire measures five aspects: mobility, self-care, daily activities, pain/discomfort, and anxiety/depression. The EQ-5D scores will be assessed at baseline and at 2, 10 and 21 weeks follow-up. In addition, cost diaries will be filled out by the patient at 10 and 21 weeks after randomization. Health care utilization including physiotherapy, visits to GP and specialists, nursing care and medication, patients costs, and absenteeism will be evaluated;

Data collection and protection

Data will be gathered using Castor EDC, a web-based secured data-capture platform (Ciwit B.V., Amsterdam, The Netherlands). This digital system will give each patient an untraceable ID. Only the main investigator, an independent monitor from the LUMC and the Inspection Healthcare and Youth (IGJ) will have access to the key that links these IDs to the patient's personal details. All data will be safeguarded and stored for 15 years after end of study. Patients will fill in digital questionnaires for all follow-up moments at home using the link they received in an e-mail. When the questionnaire is completed, it will automatically be processed in Castor EDC. Only members of the research team, an independent monitor from the LUMC and the Inspection Healthcare and Youth (IGJ) will have access to the final dataset.

Sample size

It is hypothesized that patients in the intervention group will have a mean NRS for leg pain of 4.0 on a 11-point scale at two weeks after treatment with TEI and patients in the control group will have a mean NRS for leg pain of 5.5 two weeks

after randomization. Based on literature review, a difference of 1.5 points on the NRS scale will be considered clinically relevant and a standard deviation of 2.6 is considered [12, 16]. Together with a power of 90% ($\beta = 0.10$) and a level of significance of 5% ($\alpha = 0.05$), 64 patients per group are needed. With two study groups and an estimated loss to follow-up of 10% a total of 142 patients needs to be recruited.

Statistical analysis

All study parameters will be analysed according to the intention-to-treat (ITT) principle.

Demographic statistics

Demographic data on age, gender, length, weight, tobacco exposure, alcohol abuse, history of previous radicular symptoms in the legs and usage of pain medication will be reported using mean and standard deviation or median and ranges if distribution is skewed. Presuming distribution is normal age, length and weight will be compared between groups using t-tests. Gender, tobacco exposure, alcohol abuse, history of previous radicular symptoms in the legs and usage of pain medication will be assessed using the Chi-square test.

Primary analysis

- The primary study outcome, the average score on the NRS leg pain scale at two weeks, will be compared between groups utilizing the unpaired t-tests.

Secondary analysis

- The average score on the NRS leg pain scale at 4 weeks will be compared between groups utilizing the unpaired t-tests.
- The absolute decrease of NRS leg pain, NRS back pain, ODI and increase of EQ-5D will be analysed using t-tests for the period from baseline to 2 weeks follow-up and from baseline to 4 weeks follow-up;
- Perceived recovery Likert scores will be dichotomized: 1–2 will be considered a success and 3–7 will be considered no success. Consequently, data will be compared using the Chi-square test at 2 and 4 weeks after randomization or injection to 10 and 21 weeks after randomization;
- The success data from the 2-week follow-up will be correlated to the success data at 14 weeks and 26 weeks using Chi-square tests to make predictions. In order to correlate 2-week and 4-week success data to the absolute data for NRS leg pain, NRS back pain, ODI and EQ-5D at 10 and 21 weeks, logistic regression analysis will be used;

- Baseline data will be correlated to the 2-week success data after treatment with TEI using t-tests for numerical variables and Chi-square test for categorical variables;
- The cost-effectiveness analysis will be evaluated from a healthcare perspective (the costs per extra patient with symptom relief) and a cost-utility analysis from a societal perspective (costs per Quality Adjusted Life Year (QALY)). Both analyses will be trial-based, with a time horizon of 21weeks. Use of health care and productivity will be valued according to the Dutch guidelines. QALYs will be assessed as the area under the utility curve based on the Dutch tariff for the EQ-5D. For sensitivity analysis, QALYs will be determined using the VAS for quality of life with power transformation. The average costs and patient outcome will be compared according to intention-to-treat, using net-benefit analysis, and using multiple imputation to account for missing data;

Withdrawal of subjects

Subjects can withdraw from the study at any time for any reason without consequences and will not be replaced. Subjects that refuse TEI after randomization to the TEI group, will be considered as cross-over subjects. Patients that withdraw during the study will be considered lost-to-follow-up patients. The study investigator can decide to withdraw a subject for urgent medical reasons.

Study monitoring

Data monitoring

Data monitoring will be performed by an independent monitor from the LUMC twice per year and will oversee the handling, safeguard and storage of data in accordance with the LUMC protocol and Good Research Practice (GRP) guidelines. No data monitoring committee is involved since TEI is considered usual care and risks are considered to be moderate.

Adverse event management

Adverse events (AEs) are defined as any undesirable experience occurring to a subject during the study, whether or not considered related to TEI. All adverse events reported spontaneously by the subject or observed by the investigator or his staff will be recorded. Adverse events from treatment with TEI are rare. The most common AEs reported in literature are transient exacerbation of pain (2.4%), pain at site of injection (1.1%) and accidental dural puncture (2.3%) [17].

Serious adverse events (SAEs) are defined as any untoward medical occurrence or effect that result in the subject's death or is life threatening, require hospitalisation or prolongation of existing inpatients' hospitalisation, result in persistent or significant disability or incapacity, is a congenital anomaly or birth defect or any other important medical event that did not result in any of the outcomes listed above caused by medical or surgical therapy, but could have been based upon appropriate judgement by the investigator. Elective hospitalization will not be considered as a serious adverse event. SAEs are reported by the investigator through the online web portal *ToetsingOnline* to the accredited Medical Ethics Committee (MEC) that approved the protocol within one week of first knowledge of SAEs that result in death or are life threatening followed by a period of maximum of 8 days to complete the initial preliminary report. All other SAEs will be reported within a period of maximum 15 days after first knowledge of the events.

All AEs will be followed until they have subsided or until a stable situation has been reached. Dependent of the kind of event, additional tests or medical procedures may be required for follow-up, and the subject may be referred to the general practitioner or a medical specialist.

In literature arachnoiditis and 'nervous system disorder' have been described as major complications. Despite the lack of exact data, these complications are considered to be rare [17].

Interim analysis

The investigator will annually submit a summary of the progress of the trial to the accredited MEC. Information will be provided on the date of inclusion of first subject, total number of subjects included to that point, total number of subjects that have completed the trial, serious adverse events, other problems and amendments. Additionally, the independent data monitor and the Inspection Healthcare and Youth (IGJ) will have access to interim analyses.

Premature termination of study

In case the study is ended prematurely, the sponsor will notify the accredited MEC within 15 days, including the reasons for the premature termination.

Dissemination of results

The results from the TEIAS Trial will be published in international peer reviewed neurological and neurosurgical related journals. Results will be presented at national and international spine and pain congresses. Furthermore, the results

can influence the usual care in the Netherlands and will subsequently be offered to be considered by (inter) national guideline-authors for treatment of sciatica.

DISCUSSION

The current guidelines on treatment of sciatica state that a period of conservative therapy up to 14–16 weeks is preferred and these guidelines are based on the results from the Sciatica trial by Peul et al. During this trial, 40% from the group that received continued conservative care crossed-over to the surgical group with a mean delay of 15 weeks [8]. The TEIAS-trial is focused on finding an appropriate type of conservative treatment to bridge the time during this period of conservative care in the acute phase of sciatica.

For the TEIAS study it is presumed that eligible patients will have an average duration of symptoms of 5 weeks when randomized to either the treatment arm or control arm. Therefore, follow-up at 10 weeks will correspond with the 14–16 week time point from the Sciatica trial. In the same manner, the 21 week follow-up moment will be similar to the 26 week time point in the Sciatica trial at which surgical intervention did not yield a superior patient outcome compared to conservative therapy. Therefore, the short-term effect of TEI will be evaluated at 1, 2 and 4 weeks and long-term evaluation will be performed at 10 and 21 weeks.

The burden from the pain and resulting immobility during the acute phase of sciatica have been a challenge and an adequate routine treatment strategy to alleviate symptoms during this period of 'wait-and-see' is lacking. TEI has often been proposed as an option to reduce pain and inflammation caused by mechanical compression of the nerve root and literature has demonstrated this treatment to be efficacious in part of the sciatica patients [18–21]. However, only few studies have investigated the use of TEI specifically in patients with acute sciatica (< 8 weeks). One study randomized 63 patients with a duration of symptoms between 2 and 4 weeks to either TEI or usual conservative care [11]. The TEI group experienced a greater improvement in functionality and back pain, but differences between groups were small. However, the effect of the intervention group was averaged over both responders and non-responders. Therefore, it is plausible that if the effect had been averaged only over the responders a clear effect could have been found for these patients. Moreover, the epidural injection lacked an analgesic medicament, baseline pain scores for back and leg as well as Roland Disability score were lower in the control group, and size of patient groups were small. Another study that studied TEI in the acute phase

of sciatica found that 66% of patients were considered responders within hours after an injection. The VAS score for leg pain lied around 30 mm on a 100 mm scale during the first 14 days for this group. For the non-responders the VAS leg pain score remained around 55 mm on average during these two weeks [12].

A Dutch study by Spijker-Huiges et al. investigated the effect of epidural steroid injections containing 80 mg of triamcinolone in normal saline in acute sciatica patients (2–4 weeks of symptoms) [22]. They compared this intervention in addition to usual care to usual care alone similarly to our study protocol. Results showed a small but clinically irrelevant difference between groups, but cost-utility analysis demonstrated that adding epidural steroid injections decreased total health care costs for these patients. However, despite randomization of patients baseline characteristics differed statistically significantly between groups and current anaesthesiologic guidelines advise injections with dexamethasone or Depo-Medrol instead of triamcinolone which was administered in the aforementioned study. Additionally, steroids were administered through a translaminar approach in contrast to the currently preferred transforaminal technique which will be used in our study. Furthermore, patients allocated to the treatment group received only one epidural injection. Part of these patients had a positive response to the injection but the effect deteriorated after several weeks. Therefore, additional injections might be suitable for this group of patients and in our study patients can opt for further treatment with TEI if the anaesthesiologist considers this step to be advantageous.

Due to the pragmatic nature of this study patients suffering from sciatica are directly referred by the GP to the anaesthesiologist without interference of any type of imaging. With the current organization of usual care it takes several weeks before the patient has visited the neurologist, undergone MRI examination and received TEI. In order to minimize this waiting time only a neurosurgeon will shortly examine the patient before inclusion to confirm that the patient is suffering from sciatica. This means that determination of the herniated disc level is based on medical history taking and physical examination. The neurosurgeon will advise the anaesthesiologist on the affected nerve root level, but if a disagreement exists the anaesthesiologist has the final decision on the level to be treated. In order to determine whether uncertainty with regard to the level of injection may play a role in the effectiveness of TEI the anaesthesiologist is asked to express his confidence in the correctness of the diagnosis and of the treated disc level.

An interesting feature of TEI that has only been studied before to a limited extent is its predictive capability on long-term patient outcome. In theory, if such pre-

dictive power of TEI exists, patients that will not do well on the long term based on the short-term effect of TEI can be offered surgery at an earlier time point. For instance, a patient could become eligible for surgery after two weeks based on the two-week TEI outcome and would not have to wait until the 14–16 week time point. Joswig et al. addressed this predictive feature in a prospective study [12]. The results demonstrated that a decrease of less than 50% in leg pain within one week indicated an unfavourable outcome at one month and therefore other treatment options should be considered. Furthermore, data suggested that prediction of the 3-month outcome might be possible which makes this a promising concept [13]. A study by El-Yahchouchi et al. investigated the predictive power of TEI for long-term outcome in patients with radicular pain, regardless of the presence of radiculopathy [14]. They assessed the effect of the epidural injection immediately after administration, at 2 weeks and at 2 months. Results showed that patient outcome at two months, the long-term effect, was not associated with immediate pain relief. However, the long-term effect was strongly related to pain relief at 2 weeks. They argued that the delayed effect from the anti-inflammatory steroid agent was not observable directly post-injection, but only established thereafter and thus was observable at the two-week follow-up moment. This is an indication that a predictive value might exist.

List of abbreviations

AE: adverse event; CEA: cost-effectiveness analysis; EQ-VAS: EuroQol Visual Analogue Scale; EQ-5D: EuroQoL-5 Dimension; GP: general practitioner; GRP: good research practice; IGJ: Inspection for Healthcare and Youth; ITT: intention-to-treat; LUMC: Leiden University Medical Center; MEC: Medical Ethical Committee; MRI: magnetic resonance imaging; NRS: numerical rating scale; ODI: Oswestry disability index; QALY: quality adjusted life year; RCT: randomized controlled trial; SAE: serious adverse event; SG: Spaarne Gasthuis; TEI: transforaminal epidural injection; TEIAS: transforaminal epidural injection in acute sciatica; VAS: visual analogue scale.

REFERENCES

1. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. BMJ 2007;334:1313.
2. Konstantinou K, Dunn KM. Sciatica: Review of Epidemiological Studies and Prevalence Estimates. Spine 2008;33(22):2454-72.
3. Freynhagen R, Baron R. The Evaluation of Neuropathic Components in Low Back Pain. Curr Pain Headache Rep 2009;13(3):185-190.
4. Hicks GE, Gaines JM, Shardell M, Simonsic E.M. Associations of back and leg pain with health status and functional capacity of older adults: Findings from the retirement community back pain study. Arthritis & Rheumatism 2008;59:1306–13.
5. Samanta A, Beardsley J. Sciatica: which intervention? BMJ 1999;319(7205):302-3.
6. Frymoyer JW. Back pain and sciatica. N Engl J Med 1988;318(5):291-300.
7. Lambeek LC, van Tulder MW, Swinkels ICS, Koppes LLJ, Anema JR, van Mechelen W. The Trend in Total Cost of Back Pain in the Netherlands in the Period 2002 to 2007. Spine Phila Pa 1976 2011;36(13):1050-8
8. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, et al. Surgery versus prolonged conservative treatment for sciatica. N Engl J Med 2007;356(22):2245-56.
9. Hofstede, SN, Marang-van de Mheen PJ, Wentink MM, Stiggelbout AM, Vleggeert-Lankamp CL, Vliet Vlieland TP, van Bodegom-Vos L. Barriers and facilitators to implement shared decision making in multidisciplinary sciatica care: a qualitative study. Implement Sci 2013;8(1):95
10. Hofstede SN, van Bodegom-Vos L, Wentink MM, Vleggeert-Lankamp CL, Vliet Vlieland TP, Marang-van de Mheen PJ. Most important factors for the implementation of shared decision making in sciatica care: ranking among professionals and patients. PLoS One 2014;9(4):e94176
11. Spijker-Huiges A, Winters JC, van Wijhe M, Groenier K. Steroid injections added to the usual treatment of lumbar radicular syndrome: a pragmatic randomized controlled trial in general practice. BMC Musculoskelet Disord 2014;15:341.
12. Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN. The Value of Short-Term Pain Relief in Predicting the One-Month Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg 2016;96:323-33.
13. Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN. The Value of Short-Term Pain Relief in Predicting the Long-Term Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg 2017;107:764-71.
14. El-Yahchouchi C, Wald J, Brault J, Geske J, Hagen C, Murthy N, Kaufmann T, Thielen K, Morris J, Diehn F, Amrami K, Carter R, Shelerud R, Maus T. Lumbar transforaminal epidural steroid injections: does immediate post-procedure pain response predict longer term effectiveness? Pain Med 2014;15(6):921-8.
15. Neal JM, Barrington, MJ, Brul R, Hadzic A, Hebl JR, Horlocker TT, Huntoon MA, Kopp, SL, Rathmell JP, Watson JC. The Second ASRA Practice Advisory on Neurologic Complications Associated With Regional Anesthesia and Pain Medicine: Executive Summary 2015. Reg Anesth Pain Med 2015; 40:401-430.
16. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. J Orthop Surg Res 2015;10(24):1-11

17. Section Pain and Palliative Medicine of the Dutch Anesthesiology Association (NVA). Protocol Veilig gebruik van epidurale corticosteroïden infiltraties. 2016. https://www.anesthesiologie.nl/publicaties/protocol-veilig-gebruik-epidurale-corticosteroiden-infiltraties. Accessed 8 Jan 2019.
18. Kaye AD, Manchikanti L, Abdi S et al. Efficacy of Epidural Injections in Managing Chronic Spinal Pain: A Best Evidence Synthesis. Pain Physician 2015;18(6):939-1004.
19. Bhatti AB, Kim S. Role of Epidural Injections to Prevent Surgical Intervention in Patients with Chronic Sciatica: A Systematic Review and Meta-Analysis. Cureus 2016;8(8):723.
20. Nandi J, Chowdhery A. A Randomized Controlled Clinical Trial to Determine the Effectiveness of Caudal Epidural Steroid Injection in Lumbosacral Sciatica. J Clin Diagn Res 2017;11(2):4-8.
21. Guo JR, Jin XJ, Shen HC, Wnag H, Zhou X, Liu XQ, Zhu NN. A Comparison of the Efficacy and Tolerability of the Treatments for Sciatica: A Network Meta-Analysis. Ann Pharmacother 2017;51(12):1041-1052.
22. Spijker-Huiges A, Vermeulen K, Winters JC, van Wijhe M, van der Meer K. Epidural steroids for lumbosacral radicular syndrome compared to usual care: quality of life and cost utility in general practice. Arch Phys Med Rehabil 2015;96(3):381-87.

6

Prediction of transforaminal epidural injection success in sciatica (POTEISS): a protocol for the development of a multivariable prediction model for outcome after transforaminal epidural steroid injection in patients with lumbar radicular pain due to disc herniation or stenosis

E.J.A. Verheijen, O.B.H.A.M. van Haagen, E.C. Bartels, K. van der Sloot, M.E. van den Akker-van Marle, E.W. Steyerberg, C.L.A. Vleggeert-Lankamp

BMC Neurol 2024

ABSTRACT

Background

Transforaminal epidural injections (TEI) can alleviate symptoms and help to maintain physical functioning and quality of life in patients with lumbar radicular pain. We aim to develop a prediction model for patient outcome after TEI in patients suffering from unilateral lumbar radicular pain due to lumbar disc herniation (LDH) or single-level spinal stenosis (LSS). The secondary aim is to estimate short-term patient outcome differences between LDH and LSS patients, the association between psychological variables and patient outcome, the rate of additional injections, surgery and complications, and to explore the short-term cost-effectiveness of TEI.

Methods

This study is designed as a multi-centre, observational, prospective cohort study in two large regional hospitals in the Netherlands. Patients diagnosed with unilateral lumbar radicular pain secondary to LDH or LSS and congruent with MRI findings, who are referred for TEI along usual care pathways, are eligible for study participation. A total of 388 patients with LDH or LSS will be included. A pre-defined set of demographic, clinical and radiological variables will be used as the predictors in the model. The primary outcome measure is the Numerical Rating Scale (NRS) for leg pain. Secondary outcome measures include back pain, physical functioning, perceived recovery, pain coping strategies, anxiety and depression and use of analgesics and physical therapy. Patients will be evaluated at baseline, 2 weeks and 6 weeks after treatment. NRS leg pain and Likert perceived recovery data will be used as the dependent variables in a generalized linear mixed model for prediction of TEI outcome, with internal validation of performance (explained variation) by bootstrap resampling. Cost-effectiveness for a period of 6 weeks prior to and after treatment will be performed with decision-analytic modelling.

Discussion

Patients with severe lumbar radicular pain often request additional treatment when conservative care is insufficient. TEI can offer relief of symptoms. Currently, it is not possible to predict responsiveness to this treatment for individual patients. This study is designed to explore predictors that can differentiate between patients that will and will not have a positive outcome after TEI. This information may support treatment strategies for this patient group.

BACKGROUND

Unilateral lumbar radicular pain is a common symptom of lumbar radiculopathy, a spinal condition with a reported lifetime prevalence of up to 43%, depending on the exact definition of this condition [1]. Generally, lumbar radicular pain is explained as a symptom when the function of one of the exiting lumbosacral nerve roots is affected [2]. The two most common causes are lumbar disc herniation (LDH) and degenerative lumbar spinal stenosis (LSS) due to age-related disc degeneration, hypertrophy of the ligamentum flavum and zygapophyseal joint, and osteophyte formation [3]. Both may result in radiating leg pain, which can be accompanied by back pain, and sensory and motor deficits. Lumbar radicular pain is often insufficiently explained by merely mechanical compression of the nerve root, and, therefore, it is suggested that, additionally, inflammatory and immunological processes are involved in LDH and LSS [4–6]. Disease progression, however, usually differs as patients with LDH are younger on average and have a more acute onset of symptoms which may aggravate when coughing, sneezing or leg-straightening, while patients with LSS often are older and experience gradually increasing symptoms that may exacerbate during standing or walking. Nonetheless, in both patient groups symptoms can severely limit physical functioning, decrease quality of life, and result in absenteeism. Symptoms are expected to spontaneously diminish with time in the majority of LDH patients as this condition is considered to be self-limiting. Hence, guidelines advise conservative treatment for the first weeks to months before opting for non-conservative therapies [7–9]. For LSS there is less general agreement on the most appropriate treatment strategy as symptoms usually follow a more chronic course, although some patients may experience natural resolution of symptoms as well [10–12]. For both patient groups, however, symptoms are frequently insufficiently alleviated with oral pain medication, and, due to the debilitating consequences, medical and socio-economic costs are tremendous in this patient population [13–16].

Transforaminal epidural steroid injections (TEI) have become a growingly popular treatment modality in patients with severe lumbar radicular pain to relieve symptoms. TEI aims to reduce the inflammation around the affected nerve root by limiting nerve root oedema, blocking prostaglandin synthesis and altering the conduction of nociceptive C-fibres [17–20]. The rationale behind this minimally invasive treatment is that it enables the patient to maintain physical functioning, restores quality of life, and, if satisfactorily effective, obviates the need for neurosurgical intervention. However, TEI is considered to have a temporary effect, and it remains unclear in what part of the patients symptoms may reoccur necessitating repeat injections or surgery on the short term nonetheless.

Specifically in patients with LSS, which is thought to be more chronic and less self-limiting than LDH, there is debate regarding the efficacy of TEI.

Several systematic reviews and meta-analyses have been conducted on the efficacy of TEI in patients with lumbar radicular pain due to LDH and LSS which demonstrated varying success rates ranging between 19% and 88.4% [21–30]. It is suggested that TEI is likely to be more beneficial in carefully selected subgroups of patients with lumbar radicular pain, explaining the variation and overall modest effect of this treatment in the whole patient group. Therefore, prognostic factors should be identified that could aid in improving patient selection allowing for more tailored treatment strategies in patients with lumbar radicular pain. Numerous demographic, clinical and radiological factors have been assessed for their association with TEI outcome across multiple studies, but few factors have been demonstrated to be consistently significantly associated [31–47] (Table 1). The most consistent prognostic factors for TEI included a shorter duration of symptoms in 3 out of 7 studies (OR 2.6 and 2.5 for VAS ≥ 50% and Roland-Morris disability score ≥ 40% reduction; r -0.48 and -0.36 for VAS and ODI) [38, 43], a positive Slump test in 2 out of 2 studies (Nagelkerke R^2 0.10 for VAS leg pain and 0.17 for ODI) [34], degree of nerve root compression in 3 out of 4 studies (R^2 0.06 for VAS leg pain and ODI) [35] and radiological etiology of symptoms in 2 out of 5 studies (measure of strength of association not reported). Since there was large methodological variability among studies, strength of association was not always reported and the majority of variables were evaluated in a small number of studies. A reliable predictive model suitable for clinical practice is lacking. Furthermore, psychological variables may affect patient outcome after TEI treatment similarly to spine surgery, but have not been assessed before in relation to this therapy [48, 49].

Therefore, a systematic assessment of demographic, clinical and radiological factors and their association with TEI outcome in patients with LDH or LSS is needed in a sufficiently large prospective cohort study. Since the purpose of TEI treatment is to rapidly resolve symptoms, the study's primary aim is to develop a multivariable prediction model for short-term patient outcome with internal validation of performance. In addition to this prediction model, we aim to determine the association between psychological baseline variables and patient outcome measures as this has not been explored before. Furthermore, as a secondary study objective we intend to explore the short-term cost-effectiveness of TEI. TEI is intended to relief pain symptoms, and, when the effect is satisfactory, may result in reduction of analgesic intake and physical therapy sessions. Moreover, patients with severe symptoms may take a sick-leave from work which has substantial socioeconomic costs. However, this the effect of TEI on

these variables has not been studied thoroughly before. We aim to explore the medical expenditures and absence from labour before and after TEI treatment.

Table 1 Variables associated with prediction of TEI outcome in previous studies. Only variables assessed in two studies or more were included. Variables that will be included in the core set of predictors for the model in this study are indicated by an asterisk.

Prediction factor for TEI outcome	Number of studies that found an association	References	Number of studies that did not find an association	References	Core set
Demographic factors					
Age	1	(38)	5	(31, 33, 39, 40, 45)	*
BMI	1	(38)	1	(45)	
Duration of symptoms	3	(33, 38, 43)	4	(31, 39, 40, 45)	*
Injection level	0		5	(31, 39, 41, 45, 47)	
Injection side	0		4	(31, 38, 39, 47)	
Sex	1	(38)	5	(31, 39-41, 45)	*
Clinical factors					
Pre-injection pain score	1	(38)	3	(33, 42, 45)	*
Pre-injection functionality score	1	(38)	1	(42)	
Positive Slump test	2	(34, 35)	0		
Post-injection pain score	1	(45)	1	(42)	
Radiological factors					
Lesion level	0		2	(31, 33)	
Lesion severity	1	(32)	2	(31, 36)	
Location of LDH	2	(35, 37)	2	(33, 47)	*
Degree of nerve root compression	3	(34, 35, 47)	1	(45)	*
Presence of transitional vertebrae	2	(44, 46)	1	(45)	
Radiological etiology of symptoms	2	(31, 41)	3	(33, 34, 40)	*
Type of disc degeneration	1	(35)	1	(47)	

METHODS AND DESIGN

This protocol has been aligned with the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement and checklist [50].

Design and study setting

The POTEISS (Prediction of Transforaminal Epidural Injection Success in Sciatica) study is designed as a multicentre, prospective, observational cohort study and is performed in accordance with recommendations by the PROGRESS framework [51]. The study is initiated at the Leiden University Medical Center (LUMC), Leiden, the Netherlands and enrolling patients at the outpatient pain clinic in the Spaarne Hospital (SG), Hoofddorp and Haarlem, the Netherlands and in the Groene Hart Hospital (GHZ), Gouda, the Netherlands since November 2020 with, currently, a total of 115 patients included. Patients that are referred to the pain clinic and that have a scheduled appointment for treatment with TEI as part of standard care are screened for eligibility. When meeting the inclusion criteria, patients are contacted by the study investigator and invited to participate. Follow-up will be performed by completion of a case report form at 30 min follow-up and by sending digital questionnaires through e-mail at baseline, 4 days, 2 and 6 weeks follow-up which are automatically processed in an online secure database once completed. A schematic overview of study procedures is provided in Fig. 1.

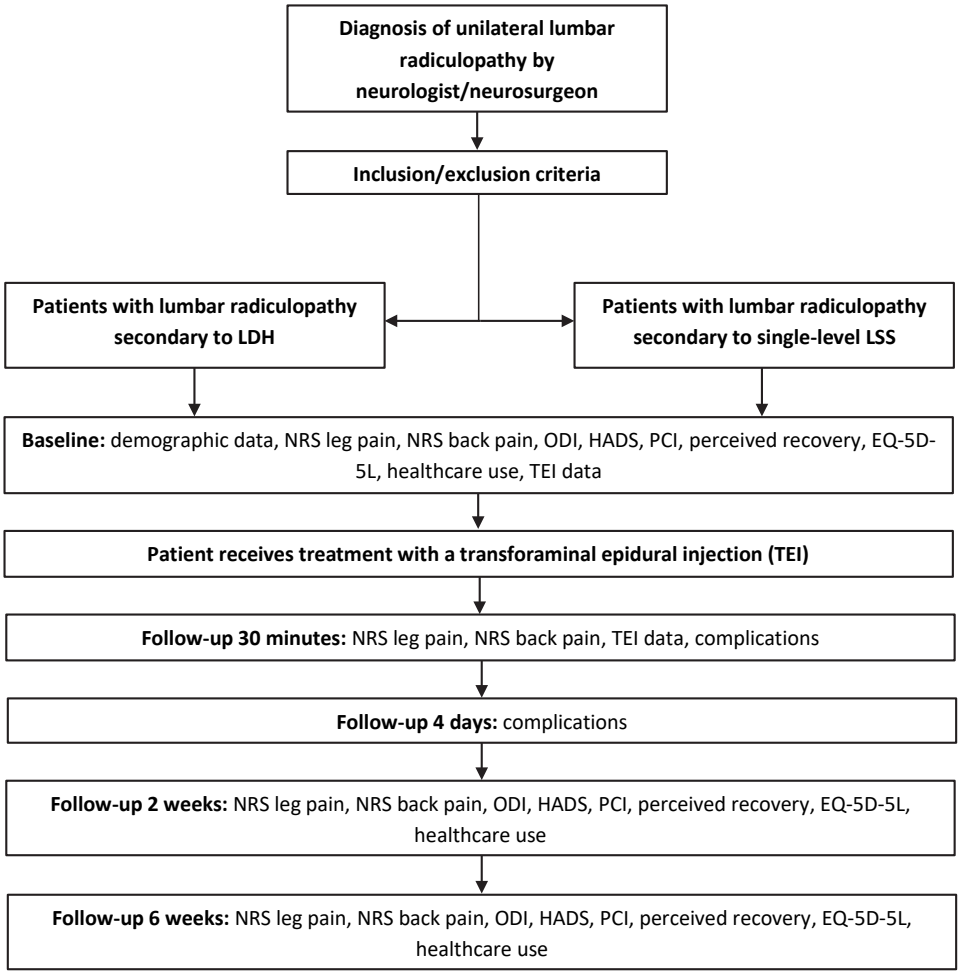


Fig. 1 Flow chart of study procedures

Aims and objectives

The aim of the POTEISS study is to explore patient outcome after TEI in patients with unilateral lumbar radicular pain secondary to lumbar disc herniation or degenerative spinal stenosis and to identify subgroups that benefit more or less from this treatment.

The primary objective is to:

1. Develop a model based on demographic, clinical and radiological variables for prediction of patient outcome after TEI (outcome defined as ≥ 30% improvement in NRS leg pain; see also *Primary analysis* section, or defined as a perceived recovery Likert score of 1 or 2; see also *Secondary analysis* section)

Secondary objectives are to:

1. Estimate short-term patient outcome after TEI in patients with LDH and LSS
2. Determine the association between psychological baseline variables and patient outcome measures
3. Determine the rate of additional injections, surgery and complications after treatment with TEI
4. Explore the short-term cost-effectiveness with decision-analytic modelling

Eligible study participants and recruitment

The patients are selected from the referral database at the pain clinic. Patients suffering from an episode of unilateral lumbar radicular pain due to LDH or single-level LSS diagnosed by the neurologist or neurosurgeon, congruent with radiological imaging (MRI), and scheduled for treatment with TEI are screened for study participation. The clinical diagnosis of lumbar radicular pain, with or without other symptoms associated with lumbar radiculopathy, by the neurologist or neurosurgeon is based on history taking and physical examination and must be supported with a radiological diagnosis of LDH or LSS on MRI. Lumbar radicular pain is defined as a painful sensation that radiates from the back towards the leg, usually going below the knee and is often described by patients as electric, shooting or cramping. Clinical presentation in case of LDH usually involves a more acute onset with a constant pain which is episodically aggravated and often not dependent on the patient's position. Patients with clinically suspected LSS usually describe a more gradual onset and progression of with unilateral radicular pain that may be aggravated by walking or standing. In both cases, MRI findings must be congruent with clinical findings, i.e., lumbar level. If MRI results are not congruent with clinical findings, the patient is not eligible for inclusion. TEI treatment appointments are made by a pain clinic nurse as part of standard care. A complete description of inclusion and exclusion criteria can be found in Table 2.

Table 2 Patient inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Clinical diagnosis of unilateral lumbar radiculopathy by neurologist or neurosurgeon	Age under 18 years Severe multisegmental spinal disease
Radiological diagnosis of LDH or LSS on MRI and congruent with clinical findings[a] Scheduled appointment for TEI Access to e-mail	Anatomical abnormalities that may complicate the procedure technically (e.g. severe scoliosis) History of lower back surgery at the same lumbar level and side

Table 2 Patient inclusion and exclusion criteria (*continued*)

Inclusion criteria	Exclusion criteria
Signed informed consent	Previous treatment with TEI for current episode of lumbar radiculopathy Active malignancy or infectious disease Use of immunosuppressive drugs Use of systemic corticosteroids in preceding 3 months
	Circumstances that prevent treatment with TEI (e.g. use of anticoagulants that cannot be temporarily discontinued, allergy against steroids or local anaesthetic)
	Pregnancy Major language barrier

[a] Definitions of LDH and LSS are provided in the *Definition of lumbar disc herniation and degenerative spinal stenosis on MRI* section. LDH = lumbar disc herniation; LSS = lumbar spinal stenosis; MRI: magnetic resonance imaging; TEI = transforaminal epidural injection

Definition of lumbar disc herniation and degenerative spinal stenosis on MRI

For this study, lumbar disc herniation is defined as the presence of protrusion, extrusion or sequestration of an intervertebral disc without other significant degenerative changes that may affect the nerve root at that level. Central lumbar disc herniation will be eligible only if symptoms are unilateral and cauda equina syndrome is absent. Patients with symptoms due to an extraforaminal disc herniation are included as well.

Degenerative spinal stenosis is defined as narrowing of the spinal canal, lateral recess or neuroforamen due to degenerative changes of the bony structures (osteophytes) and surrounding tissues including ligaments, zygapophyseal joints, and a bulging intervertebral disc. Only patients with a stenosis causing unilateral symptoms are eligible. Patients with synovial (facet) cysts, neoplasm or spondylolisthesis resulting in stenosis will be excluded.

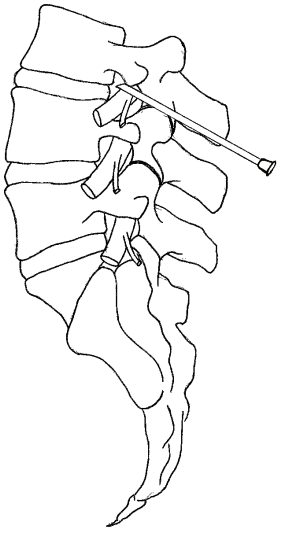
Evident physical compression of the nerve root will not be an eligibility criterion, but compression will be graded for analysis of MRI data (see also *Secondary study parameters* section), since in patients without evident nerve root compression inflammatory processes may be a key contributor that can be treated with a transforaminal epidural steroid injection [52–54].

Intervention procedures

Transforaminal epidural injection (TEI)

Administration of the transforaminal epidural injection (TEI) will be performed by multiple experienced anaesthesiologists from the participating hospitals. The treating anaesthesiologist determines the treatment level based on clinical findings and radiological imaging (diagnosed by neurologist or neurosurgeon). The patient will lie in a prone position on the table with a support underneath to reduce the natural lordosis. The skin is sterilized with chlorhexidine and locally anaesthetized with 1 mL 1% lidocaine. Under fluoroscopic image guidance the needle is placed via a transforaminal approach in close proximity of the nerve root. This is achieved by advancing the needle subpedicularly towards the safe triangle which is confined by the inferior border of the pedicle, the exiting nerve root and inferiorly from the anterior margin of the pedicle [55] (Fig. 2).

Fig. 2 Transforaminal approach into the safe triangle space. This space is confined by the inferior border of the pedicle, the exiting nerve root and inferiorly from the anterior margin of the pedicle



Contrast agent is injected to confirm correct positioning of the needle. In case of incorrect positioning, the needle is retracted and repositioned towards the correct location. In accordance with current Dutch anaesthesiologic guidelines, injections at L3 and below will contain 1,5 ml lidocaine 2% and 40 mg methylprednisolone acetate and injections above L3 will contain 1,5 ml lidocaine 1% and 10 mg dexamethasone to prevent possible vascular occlusion of the artery

of Adamkiewicz [56]. After the procedure, the patient will stay for 30 min at the recovery room for monitoring.

Adjuvant care

Since this is an observational cohort study, patients are allowed to use any type of adjuvant care as they require. This may include, but is not limited to, analgesic usage and physical therapy. To compare the use of concomitant therapy between groups, the patients are requested to register their use of oral pain medication and physical (rehabilitation) therapy during follow-up.

Data collection and protection

Data will be collected using Castor EDC, a web-based data capture platform (Ciwit B.V., Amsterdam, The Netherlands). After study participation has been confirmed and informed consent has been provided, the patient will complete baseline questionnaires the day before TEI treatment. For all follow-up moments the patient will receive an e-mail with a hyperlink to the questionnaires to fill in at home. When the questionnaires are completed, they will be automatically processed anonymously in Castor EDC. All data will be safeguarded and stored for 15 years after end of study. In the database each patient is given an untraceable ID that can only be linked to the patient's personal details with a key file. Access to this key file and the final dataset will only be granted to members of the research team, an independent monitor and the Inspection for Healthcare Inspection and Youth (IGJ).

Study parameters and follow-up

Demographic data are assessed at baseline by an independent researcher and outcome measurements are self-reported by the patient. Follow-up starts on the day of treatment and patients fill in questionnaires at 4 days, 2 and 6 weeks after treatment. Patients are not entitled to see results from previous follow-up moments. A summary of all outcome measurements for each follow-up moment can be found in Table 3.

Table 3 Data collection schedule

	Baseline	Follow-up			
		30 min.	Day 4	Week 2	Week 6
Demographic data	✓				
NRS leg pain	✓	✓		✓	✓
NRS back pain	✓	✓		✓	✓
ODI	✓			✓	✓
HADS	✓			✓	✓
PCI	✓			✓	✓
Perceived recovery				✓	✓
EQ-5D-5L	✓			✓	✓
Health care use[a]	✓			✓	✓
TEI data[a]		✓	✓		
Complications		✓	✓		

[a] Specifics on the collected data are provided in the *Secondary study parameters* section. HADS = hospital anxiety and depression scale; NRS = numerical rating scale; ODI = Oswestry disability index; PCI = pain coping inventory; TEI = transforaminal epidural injection; VAS = visual analogue scale

Primary study parameter

- The primary study outcome measure, the Numerical Rating Scale (NRS) for leg pain, will be assessed at baseline, 30 min, 2 weeks and 6 weeks after treatment. NRS for leg pain will be scored on a 0 (no pain) to 10-point (worst imaginable pain) scale with increments of 1 point. Each patient will give an average score for the previous 7 days and the highest score during those same 7 days. Leg pain from lumbar radiculopathy may vary during the week and, therefore, an average score of the past 7 days will be more representative with sufficient reliability [57, 58].

Secondary study parameters

In order to assess back pain, physical functioning, anxiety and depression-related complaints, pain coping strategies and perceived recovery secondary study outcome measures will be measured. In addition, other study parameters will be collected for development of a prediction model.

- NRS for back pain will be scored at baseline, 30 min, 2 weeks and 6 weeks after treatment. NRS for back pain will be scored on a 0 (no pain) to 10-point (worst imaginable pain) scale with increments of 1 point. Each patient will give an average score for the previous 7 days and the highest score during those same 7 days. Similarly to leg pain, back pain intensity may change dur-

ing the week and, therefore, an average score of the past 7 days will be more representative with sufficient reliability [57, 58].

- The Oswestry Disability Index (ODI) will be recorded to assess physical functioning. The ODI will be scored on a 0 (no disability) to 50-point (worst disability possible) scale at baseline, 2 and 6 weeks after treatment and multiplied by a factor of 2 to provide a score on a 0 to 100 percentage scale.
- The Hospital Anxiety and Depression Scale (HADS) will be measured to evaluate complaints of anxiety and depression at baseline, 2 weeks and 6 weeks after treatment. Both are scored with 7 items on a Likert scale that ranges from 0 to 3. For each dimension a maximum score of 21 points is possible. A score between 0 and 7 implies the absence of anxiety or depression, between 8 and 10 indicates a possible anxiety or depression disorder, and 11 or higher indicates the presence of anxiety or depression.
- The Pain Coping Inventory (PCI) will be recorded to measure how the patient copes with pain symptoms. This scale assesses 6 different coping styles using 33 items which can be answered with scores from 1 (never or rarely used) to 4 (very often used) to determine how often the patient practices a certain coping strategy. The PCI will be obtained at baseline, 2 weeks and 6 weeks after treatment.
- Perceived recovery will be measured to determine how the patient perceived his or her recovery. A 7-point Likert scale is used: 'completely recovered' = 1, 'significantly recovered' = 2, 'somewhat recovered' = 3, 'no recovery' = 4, 'somewhat worsened' = 5, 'significantly worsened' = 6 and 'worse than ever' = 7. Perceived recovery will be assessed at 2 weeks and 6 weeks after treatment.
- EQ-5D-5L questionnaire: will be used for the cost-effectiveness analysis. This questionnaire consists of both the EQ-5D-5L descriptive systems which measures five dimensions of health-related Quality of Life (mobility, self-care, daily activities, pain/discomfort, and anxiety/depression with 5 severity levels) and the EQ-VAS, a visual analogue scale rating the overall current health-related Quality of Life ranging from 0 (as bad as death) to 100 (perfect health). The EQ-5D scores will be assessed at baseline and at 2, and 6 weeks follow-up.
- Health care use: pain medication use, physical therapy use and number of appointments with GP (at GP's practice, at home or via telephone/e-mail) are recorded at 2 weeks and 6 weeks after treatment. At baseline patients are asked about these variables for the 6-week period preceding study participation.
- Radiological variables: lumbar level and location of disc herniation (central, paramedian, foraminal or extraforaminal) or stenosis (central, lateral recess, foraminal), type of disc degeneration (bulging, protrusion, extrusion, seques-

tration) and Pfirrmann classification [59], degree of stenosis (Miskin scale [60]), disc height loss (0%, 1–25%, 26–50%, 51–75%, >75%), vertebral end plate changes (Modic scale [61]), facet degeneration (Weishaupt scale [62]), epidural lipomatosis grading (Borré scale [63]) and degree of nerve root impingement (Pfirrmann grading [64]) will be evaluated by two independent researchers, blinded to clinical data including patient outcome, on T1 and T2 weighted sagittal and transversal MRI data.

Other parameters

- TEI data: data on administration of TEI is collected by the anaesthesiologist or by a nurse. This includes the date of injection, lumbar level, side and the occurrence of any complication during the procedure or within 30 min afterwards.
- Data regarding the need for additional injections or surgery after TEI will be obtained from the patient's medical record.
- Complications: data will be collected on complications occurring periprocedural by the anaesthesiologist or nurse. At the fourth day after treatment the patient is requested to fill in a questionnaire about the occurrence of any complication since the procedure. Information on other adverse events that require a visit to the hospital or hospitalization will be obtained from the patient's medical record.

Sample size calculation

The sample size was calculated for a multivariable prediction model with a binary outcome in R (version 3.6.3) [65] using the 'pmsampsize' package [66] to ensure ample patients for this analysis. The criteria used to calculate the sample size aim to minimize overfitting and ensure accurate estimation of pivotal variables in the model [67]. For the calculation we used a dichotomized NRS for leg pain as the binary outcome variable based on a minimal reduction of 30% in leg pain intensity compared to baseline (further elaborated in the 'Analysis' section). We hypothesized that 60% of the patients will be considered a 'success' (for definition see *Primary analysis* section) after two weeks based on estimations in previous literature [25]. We assumed an R^2 (percentage of variation in outcome explained by the model) of 0.25, based on the strength of association of individual predictors as reported by a limited number of previous studies. Furthermore, we assumed an expected number of predictor variables of 10 and a desired level of shrinkage (measure of overfitting) at internal validation of 0.9. With an assumed margin of error of 0.05 in the estimation of the intercept a total of 369 patients is needed. A sample size of 369 patients will suffice for a range of R^2 (0.22–0.73). With a loss to follow-up of 5% we aim to

include a total of 388 patients. The ratio of patients with LDH and LSS to be included will not be predefined but will result from the ratio in clinical practice. We estimated that the ratio LDH: LSS patients will be between 2:1 and 3:1 based on clinical experience.

Data analysis

Demographic data will be reported using descriptive statistics. Continuous variables (age, BMI, duration of symptoms) will be presented by median and interquartile range. Categorical variables (sex, cause of symptoms, smoking status, alcohol use, side of radicular symptoms, use of pain medication, use of physical therapy, living situation and daily occupation) will be presented using numbers and percentages. For comparisons of demographic data between patients with LDH and LSS, appropriate statistical tests will be performed, depending on scale of measurement: for continuous outcomes the Mann–Whitney U test will be used, whereas for categorical data the χ^2 test will be employed.

Primary analysis

The primary study outcome measure, the average score on the NRS leg pain scale, will be determined 30 min after the procedure and at the 2 and 6-week follow-up moments. Linear mixed models will be used with a radiological diagnosis (LDH or LSS) by time (as a main effect) and random effect for repeated measures within-subjects to assess differences between patient groups across all time points while correcting for potential confounders.

Furthermore, for each patient the change in NRS leg pain score at follow-up will be dichotomized using a predefined cut-off of 30% or more pain reduction compared to baseline. This is based on the minimally clinical important difference (MCID) of 30% as recommended by an international consensus group [68]. Patients with pain reduction of 30% or more compared to baseline will be considered 'success' and patients with less pain reduction, an increase in pain, repeat injections or surgery within the follow-up period will be considered 'non-success'. Proportions of 'success' and 'non-success' will be compared between groups across all time-points. A generalized linear mixed model analysis for this dichotomized outcome measure at 6 weeks follow-up will be performed to predict a positive outcome after TEI. The model will include different sets of predictors using demographic, clinical and radiological variables:

- Core predictors: age, sex, duration of symptoms, pre-injection NRS for leg pain, HADS and PCI scores, radiological diagnosis (LDH or single-level LSS), location of disc degeneration, degree of nerve root compression (Pfirrmann grading) and vertebral endplate changes (Modic scale).

- Extended set of predictors: body mass index (BMI), injection level, injection side, pre-injection ODI score, presence of neurologic deficit (sensory and motor deficits observed during physical examination), post-procedural NRS for leg pain score at 30 min, NRS for leg pain at 2 weeks, perceived recovery at 2 weeks, type of disc degeneration, MRI signal intensity of the nerve root, lesion level and presence of transitional vertebrae
- Exploration of MRI: radiological variables not included in the core or extended predictor set will be assessed as categorical variables using different scales (see *Secondary study parameters* section) for their association with TEI outcome

The core predictor set includes variables that were relatively consistently associated with TEI outcome based on a literature review and variables that were hypothetically auspicious.

Performance measures will include the area under the receiver operating characteristic curve (AUC) to indicate discriminative ability, and R^2 for overall performance. Internal validation will be performed by bootstrap resampling with 500 iterations.

Secondary analysis

The mean NRS for back pain and ODI will be determined postprocedural and for the 2 and 6-week follow-up moments. Similarly to NRS for leg pain, linear mixed models will be used with a radiological diagnosis (LDH or LSS) by time (as a main effect) and random effect for repeated measures within-subjects to assess differences between patient groups across all time points while correcting for potential confounders. For assessment of the number of 'success' and 'non-success' for back pain (NRS) and disability (ODI) scores, a predefined cut-off of 30% or more reduction compared to baseline will be used [68].

Likert data for perceived recovery will be dichotomized (a score of 1 or 2 will be considered a 'successful recovery' and patients with this score, therefore, 'success', 3–7 will be considered 'no successful recovery' and patients with this score, therefore, 'non-success') and compared between LDH and LSS groups using Chi-square tests for the 2- and 6-week follow-up time points. A generalized linear mixed model analysis for dichotomized perceived recovery data at 6-weeks follow-up will be performed. The same sets of predictors will be used as for the prediction model based on dichotomized NRS leg pain data.

An exploratory cost-effectiveness analysis of TEI versus usual care will be performed by comparing the observed costs and effects the six weeks before treatment to the six weeks after treatment, assuming the period before treatment to be representative for usual care. The analysis will be performed from a healthcare perspective. Patients will provide data for this analysis through the EQ-5D-5L and healthcare use questionnaires at baseline, 2 and 6 weeks follow-up. Healthcare use will be valued using standard prices as published in the Dutch costing guideline [69]. Effects will be expressed in quality adjusted life years (QALYs). QALYs will be estimated by first calculating the utilities using the Dutch tariff [70]. Subsequently, the QALYs for the 6 week period before treatment will be based on the utility at baseline, and the QALYs in the 6 weeks after treatment will be obtained from utilities measured at baseline, 2 and 6 weeks follow-up using the area under the curve method.

Missing data

Patterns of missing data will be explored and visualized. Imputation will be considered using multiple imputation based on correlation with covariates and the outcomes using the 'mice' package in R [71].

Reporting

Study results will be reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines [50, 72].

Study limitations

This study is limited by the absence of a formal control group. It is possible that a subgroup of patients with lumbar radicular pain can be satisfactorily treated with an inert substance or other types of treatment [27]. As such, it is possible that prognostic factors found through this study may also be associated with outcome after other treatments. Ideally, this would require a randomized controlled trial and the sample size would have to be smaller for feasibility reasons precluding the development of a prediction model for treatment outcome. The exploratory cost-effectiveness assumes the 6-week period before treatment to be representative of usual care. Additionally, the study may be limited by selection bias since most patients will be suffering from chronic symptoms. Patients with (sub)acute symptoms are usually treated conservatively. At the time they are referred to the pain clinic for TEI, symptoms will have a more chronic character. Furthermore, at baseline patients are asked about their health care utilization in the six weeks prior to participation. This may possibly lead to recall bias.

DISCUSSION

Patients with severe lumbar radicular pain often request further treatment when conservative care with oral pain medication is inadequate. TEI is a minimally invasive treatment that could provide sufficient symptom reduction and obviate the need for surgery in part of those patients for whom conservative care have failed. To offer each patient the most appropriate treatment strategy it is necessary that clinicians can differentiate between patients whose symptoms can be sufficiently alleviated with TEI and patients that may need to be subjected to surgery nonetheless. Therefore, this study aims to explore prognostic factors and develop a prediction model for assessment of the likely outcomes with TEI.

A considerable number of studies has previously assessed prognostic factors for TEI success in patients with lumbar radicular pain. However, only few variables have been evaluated across multiple studies and for these factors the results are inconsistent. Among demographic variables, duration of symptoms has been demonstrated to be associated with pain reduction across multiple studies. A prolonged duration of symptoms was associated with worse outcome in three studies [33, 38, 43]. Yet, in the majority of studies duration of symptoms along with age, BMI, sex, injection level and side were not significantly associated with outcomes. Additionally, among clinical variables a positive Slump test was shown to be associated with patient outcome in two studies [34, 35]. Pre-injection pain and physical functioning scores, and post-injection pain scores were not associated with TEI outcome in multiple studies. Furthermore, several radiological characteristics were correlated with treatment outcome. Two studies showed an association between superior TEI outcome and LSS rather than LDH [31, 41], although three studies reported no association [33, 34, 40], and two studies reported that patients with transitional vertebrae fared worse after TEI compared to patients without this spinal deformity [44, 46]. Four studies that assessed the association between the location of disc herniation and treatment outcome showed inconsistent results [33, 34, 37, 47]. In addition, three out of four studies demonstrated better outcome after TEI in patients with a higher degree of nerve root impingement [34, 35, 47]. This appears paradoxical as it is hypothesized that in patients without high-grade nerve root compression inflammation plays a more crucial role and, expectingly, would fare better after TEI treatment. Therefore, the degree of nerve root compression will be thoroughly evaluated in the current study and included as a covariate in the prediction model. Additionally, the HADS and PCI scores will be included in the prediction model. The PCI has not been correlated with outcome after TEI before, while the HADS has been investigated before only once, but that study included patients undergoing multiple injections [73]. However, both have been demonstrated to

correlate with outcome after surgery predicting residual pain and disability and a passive pain coping strategy may aid in the transition from acute to chronic pain [74–77]. Furthermore, among radiological variables the presence of Modic changes will be included as it is hypothesized that this feature represents the effects of inflammatory processes which may respond better to treatment with steroids [78–80].

The large discrepancies between findings from previous studies most likely relate to the variety of patient populations, methodological designs and treatment techniques. For this reason, it is imperative to provide a clear definition of patients with lumbar disc herniation, but particularly of stenosis patients. Patients with central disc herniation or spinal stenosis and radicular pain in both legs may require treatment with bilateral transforaminal injections. Moreover, in patients with symptomatic multi-level spinal stenosis a corticosteroid injection at a single level may not be sufficient and, therefore, outcome measures will not accurately present the effect of the injection. Hence, in the current study only patients are included with unilateral radicular pain congruent with radiological imaging and without clinically relevant multi-level degenerative changes. Although TEI may be effective for patients suffering from unilateral lumbar radicular pain due to LDH as well as single-level LSS, some variables could predict treatment outcome in one group but not the other. A number of previous studies has evaluated both aetiologies as a single population and possibly this obscured the true predictive value of certain factors. Therefore, in this study these patient groups will be assessed separately.

Additionally, variation in methodological design may have affected the reported effect sizes of TEI in previous studies. The sample sizes of some studies were small, which may have led to a lack of statistical power to draw appropriate conclusions. For this prospective study, the sample size calculation was specifically aimed at the development of a prediction model to ensure sufficient power for detection of significant differences. Moreover, treatment outcome after TEI is often presented in terms of success rates based on a predefined percentual or absolute decline in pain or disability scores. However, various definitions of success were employed which reduces comparability between studies. We will use a standardized threshold of 30% pain reduction compared to baseline as recommended by an international consensus group [68], although, statistically, we recognize a number of limitations of dichotomization, such as debate on whether to use an absolute of relative scale, the arbitrary point of dichotomization and a loss of statistical power.

Finally, treatment procedures were different among previous studies. The diversity of corticosteroids, injection dose and number of injections may have affected treatment results. In the current study, standardized doses of corticosteroids will be used according to national anesthesiologic guidelines and patients will only receive one injection. In patients receiving a second injection within six weeks after the first injection, initial treatment will be considered to have failed [48].

This observational cohort study will provide insights in the outcome after a single TEI in patients with unilateral lumbar radicular pain secondary to LDH or single-level LSS in clinical practice. In addition, this study is expected to provide data for the development of a prediction model of patient outcome after TEI. Prediction of patient outcome after TEI could be an important step towards a tailored and personalized treatment strategy for patients with lumbar radicular pain.

List of abbreviations

AUC: area under the curve; BMI: body mass index; EQ-VAS: EuroQoL visual analogue scale; GHZ: Groene Hart Hospital; GP: general practitioner; HADS: hospital anxiety and depression scale; IGJ: Inspection for Healthcare and Youth; LDH: lumbar disc herniation; LSS: single-level lumbar spinal stenosis; LUMC: Leiden University Medical Center; MCID: minimally clinical important difference; MRI: magnetic resonance imaging; NRS: numerical rating scale; ODI: Oswestry disability index; OR: odds ratio; PCI: pain coping inventory; POTEISS: Prediction of Transforaminal Epidural Injection Success in Sciatica; QALY: quality-adjusted life year; SG: Spaarne Hospital; TEI: transforaminal epidural injection; VAS: visual analogue scale

REFERENCES

1. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. Spine (Phila Pa 1976). 2008;33(22):2464-72.
2. Tarulli AW, Raynor EM. Lumbosacral radiculopathy. Neurol Clin. 2007;25(2):387-405.
3. Waldman SD. 84 - Spinal Stenosis. In: Waldman SD, editor. Atlas of Common Pain Syndromes (Fourth Edition). Philadelphia: Elsevier; 2019. p. 324-7.
4. Kobayashi S, Yoshizawa H, Yamada S. Pathology of lumbar nerve root compression. Part 1: Intraradicular inflammatory changes induced by mechanical compression. J Orthop Res. 2004;22(1):170-9.
5. Robinson JP, Brown PB, Fisk JD. Pathophysiology Of Lumbar Radiculopathies and The Pharmacology Of Epidural Corticosteroids and Local Anesthetics. Physical Medicine and Rehabilitation Clinics of North America. 1995;6(4):671-90.
6. Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth. 2007;99(4):461-73.
7. Chiu CC, Chuang TY, Chang KH, Wu CH, Lin PW, Hsu WY. The probability of spontaneous regression of lumbar herniated disc: a systematic review. Clin Rehabil. 2015;29(2):184-95.
8. Kreiner DS, Hwang SW, Easa JE, Resnick DK, Baisden JL, Bess S, et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. Spine J. 2014;14(1):180-91.
9. Lorio M, Kim C, Araghi A, Inzana J, Yue JJ. International Society for the Advancement of Spine Surgery Policy 2019-Surgical Treatment of Lumbar Disc Herniation with Radiculopathy. Int J Spine Surg. 2020;14(1):1-17.
10. Kreiner DS, Shaffer WO, Baisden JL, Gilbert TJ, Summers JT, Toton JF, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). Spine J. 2013;13(7):734-43.
11. Anderson DB, Luca K, Jensen RK, Eyles JP, Van Gelder JM, Friedly JL, et al. A critical appraisal of clinical practice guidelines for the treatment of lumbar spinal stenosis. Spine J. 2021;21(3):455-64.
12. Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and Management of Lumbar Spinal Stenosis: A Review. Jama. 2022;327(17):1688-99.
13. Geurts JW, Willems PC, Kallewaard JW, van Kleef M, Dirksen C. The Impact of Chronic Discogenic Low Back Pain: Costs and Patients' Burden. Pain Res Manag. 2018;2018:4696180.
14. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J. 2008;8(1):8-20.
15. Lambeek LC, van Tulder MW, Swinkels IC, Koppes LL, Anema JR, van Mechelen W. The trend in total cost of back pain in The Netherlands in the period 2002 to 2007. Spine (Phila Pa 1976). 2011;36(13):1050-8.
16. Gore M, Sadosky A, Stacey BR, Tai KS, Leslie D. The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings. Spine (Phila Pa 1976). 2012;37(11):E668-77.
17. McLain RF, Kapural L, Mekhail NA. Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J. 2005;5(2):191-201.

18. Collighan N, Gupta S. Epidural steroids. Continuing Education in Anaesthesia Critical Care & Pain. 2009;10(1):1-5.
19. Johansson A, Hao J, Sjölund B. Local corticosteroid application blocks transmission in normal nociceptive C-fibres. Acta Anaesthesiol Scand. 1990;34(5):335-8.
20. Kantrowitz F, Robinson DR, McGuire MB, Levine L. Corticosteroids inhibit prostaglandin production by rheumatiod synovia. Nature. 1975;258(5537):737-9.
21. Manchikanti L, Buenaventura RM, Manchikanti KN, Ruan X, Gupta S, Smith HS, et al. Effectiveness of therapeutic lumbar transforaminal epidural steroid injections in managing lumbar spinal pain. Pain Physician. 2012;15(3):E199-245.
22. Buenaventura RM, Datta S, Abdi S, Smith HS. Systematic review of therapeutic lumbar transforaminal epidural steroid injections. Pain Physician. 2009;12(1):233-51.
23. Benny B, Azari P. The efficacy of lumbosacral transforaminal epidural steroid injections: a comprehensive literature review. J Back Musculoskelet Rehabil. 2011;24(2):67-76.
24. Bhatia A, Flamer D, Shah PS, Cohen SP. Transforaminal Epidural Steroid Injections for Treating Lumbosacral Radicular Pain from Herniated Intervertebral Discs: A Systematic Review and Meta-Analysis. Anesth Analg. 2016;122(3):857-70.
25. MacVicar J, King W, Landers MH, Bogduk N. The effectiveness of lumbar transforaminal injection of steroids: a comprehensive review with systematic analysis of the published data. Pain Med. 2013;14(1):14-28.
26. Oliveira CB, Maher CG, Ferreira ML, Hancock MJ, Oliveira VC, McLachlan AJ, et al. Epidural Corticosteroid Injections for Sciatica: An Abridged Cochrane Systematic Review and Meta-Analysis. Spine (Phila Pa 1976). 2020;45(21):E1405-e15.
27. Verheijen EJA, Bonke CA, Amorij EMJ, Vleggeert-Lankamp CLA. Epidural steroid compared to placebo injection in sciatica: a systematic review and meta-analysis. Eur Spine J. 2021.
28. Chou R, Hashimoto R, Friedly J, Fu R, Bougatsos C, Dana T, et al. Epidural Corticosteroid Injections for Radiculopathy and Spinal Stenosis: A Systematic Review and Meta-analysis. Ann Intern Med. 2015;163(5):373-81.
29. Manchikanti L, Knezevic NN, Boswell MV, Kaye AD, Hirsch JA. Epidural Injections for Lumbar Radiculopathy and Spinal Stenosis: A Comparative Systematic Review and Meta-Analysis. Pain Physician. 2016;19(3):E365-410.
30. Manchikanti L, Knezevic E, Latchaw RE, Knezevic NN, Abdi S, Sanapati MR, et al. Comparative Systematic Review and Meta-Analysis of Cochrane Review of Epidural Injections for Lumbar Radiculopathy or Sciatica. Pain Physician. 2022;25(7):E889-e916.
31. Bang YJ, Park HJ, Sim WS, Lee DW, Lee JY. Correlation between paravertebral spread of injectate and clinical efficacy in lumbar transforaminal block. Sci Rep. 2020;10(1):11508.
32. Chang MC, Lee DG. Outcome of Transforaminal Epidural Steroid Injection According to the Severity of Lumbar Foraminal Spinal Stenosis. Pain Physician. 2018;21(1):67-72.
33. Cyteval C, Fescquet N, Thomas E, Decoux E, Blotman F, Taourel P. Predictive factors of efficacy of periradicular corticosteroid injections for lumbar radiculopathy. AJNR Am J Neuroradiol. 2006;27(5):978-82.
34. Ekedahl H, Jönsson B, Annertz M, Frobell RB. Three week results of transforaminal epidural steroid injection in patients with chronic unilateral low back related leg

pain: The relation to MRI findings and clinical features. J Back Musculoskelet Rehabil. 2016;29(4):693-702.
35. Ekedahl H, Jönsson B, Annertz M, Frobell RB. The 1-Year Results of Lumbar Transforaminal Epidural Steroid Injection in Patients with Chronic Unilateral Radicular Pain: The Relation to MRI Findings and Clinical Features. Am J Phys Med Rehabil. 2017;96(9):654-62.
36. Elsayad E, Mortada M, Elhawala TAES, Hady EAE, Hijab A. SAT0596 Efficacy of epidural steroid injection in lumbar spinal stenosis is not related to the degree of severity by mri. Annals of the Rheumatic Diseases. 2017;76(Suppl 2):1000-.
37. Guclu B, Deniz L, Yuce Y, Adilay U, Aytar H, Turkoglu M, et al. Transforaminal Epidural Steroid Injection in the Treatment of Pain in Foraminal and Paramedian Lumbar Disc Herniations. Turk Neurosurg. 2020;30(3):394-9.
38. Kaufmann TJ, Geske JR, Murthy NS, Thielen KR, Morris JM, Wald JT, et al. Clinical effectiveness of single lumbar transforaminal epidural steroid injections. Pain Med. 2013;14(8):1126-33.
39. Kim YH, Kim HJ, Seo J, Chai JW, Song HG, Choi YH, et al. Spinal nerve signal intensity on Dixon T2-weighted water-only sequence: an important outcome predictor after lumbar transforaminal epidural injection. Eur Radiol. 2021;31(12):9459-67.
40. Lee JW, Kim SH, Lee IS, Choi JA, Choi JY, Hong SH, et al. Therapeutic effect and outcome predictors of sciatica treated using transforaminal epidural steroid injection. AJR Am J Roentgenol. 2006;187(6):1427-31.
41. Olguner SK, Celiktas M, Oktay K, Arslan A, Bilgin E, Gezercan Y, et al. Comparison of 1-year results of single transforaminal epidural steroid injection among patients with different spinal pathologies-related radicular pain. Niger J Clin Pract. 2020;23(6):835-41.
42. Ploumis A, Christodoulou P, Wood KB, Varvarousis D, Sarni JL, Beris A. Caudal vs transforaminal epidural steroid injections as short-term (6 months) pain relief in lumbar spinal stenosis patients with sciatica. Pain Med. 2014;15(3):379-85.
43. Sariyildiz MA, Batmaz İ, Yazmalar L, Güneş M, Turan Y. The effectiveness of transforaminal epidural steroid injections on radicular pain, functionality, psychological status and sleep quality in patients with lumbar disc herniation. J Back Musculoskelet Rehabil. 2017;30(2):265-70.
44. Sencan S, Azizov S, Celenlioglu AE, Bilim S, Gunduz OH. Effect of sacralization on the success of lumbar transforaminal epidural steroid injection treatment: prospective clinical trial. Skeletal Radiol. 2022.
45. Şencan S, Çelenlioğlu AE, Asadov R, Gündüz OH. Predictive factors for treatment success of transforaminal epidural steroid injection in lumbar disc herniation-induced sciatica. Turk J Med Sci. 2020;50(1):126-31.
46. Son KM, Lee SM, Lee GW, Ahn MH, Son JH. The Impact of Lumbosacral Transitional Vertebrae on Therapeutic Outcomes of Transforaminal Epidural Injection in Patients with Lumbar Disc Herniation. Pain Pract. 2016;16(6):688-95.
47. Tecer D, Adiguzel E, Tan AK, Taskaynatan MA. Role of Magnetic Resonance Imaging in Ascertaining the Success of Transforaminal Epidural Steroid Injection for Lumbar Radicular Pain. Pain Med. 2017;18(4):645-50.
48. Johansson AC, Linton SJ, Rosenblad A, Bergkvist L, Nilsson O. A prospective study of cognitive behavioural factors as predictors of pain, disability and quality of life one year after lumbar disc surgery. Disabil Rehabil. 2010;32(7):521-9.

49. Ionescu D, Iacob CI, Brehar FM, Avram E. The role of catastrophizing and basic psychological needs satisfaction on health-related quality of life and pain in patients with lumbar disc herniation. Front Psychol. 2023;14:1147254.
50. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63.
51. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. Bmj. 2013;346:e5793.
52. el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, Van der Kallen BF, van den Hout WB, Jacobs WC, et al. Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med. 2013;368(11):999-1007.
53. Janardhana AP, Rajagopal, Rao S, Kamath A. Correlation between clinical features and magnetic resonance imaging findings in lumbar disc prolapse. Indian J Orthop. 2010;44(3):263-9.
54. van Rijn JC, Klemetso N, Reitsma JB, Majoie CB, Hulsmans FJ, Peul WC, et al. Symptomatic and asymptomatic abnormalities in patients with lumbosacral radicular syndrome: Clinical examination compared with MRI. Clin Neurol Neurosurg. 2006;108(6):553-7.
55. Mandell JC, Czuczman GJ, Gaviola GC, Ghazikhanian V, Cho CH. The Lumbar Neural Foramen and Transforaminal Epidural Steroid Injections: An Anatomic Review With Key Safety Considerations in Planning the Percutaneous Approach. AJR Am J Roentgenol. 2017;209(1):W26-w35.
56. Dietrich TJ, Sutter R, Froehlich JM, Pfirrmann CW. Particulate versus non-particulate steroids for lumbar transforaminal or interlaminar epidural steroid injections: an update. Skeletal Radiol. 2015;44(2):149-55.
57. Bolton JE, Humphreys BK, van Hedel HJ. Validity of weekly recall ratings of average pain intensity in neck pain patients. J Manipulative Physiol Ther. 2010;33(8):612-7.
58. Jamison RN, Raymond SA, Slawsby EA, McHugo GJ, Baird JC. Pain assessment in patients with low back pain: comparison of weekly recall and momentary electronic data. J Pain. 2006;7(3):192-9.
59. Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine (Phila Pa 1976). 2001;26(17):1873-8.
60. Miskin N, Isaac Z, Lu Y, Makhni MC, Sarno DL, Smith TR, et al. Simplified Universal Grading of Lumbar Spine MRI Degenerative Findings: Inter-Reader Agreement of Non-Radiologist Spine Experts. Pain Med. 2021;22(7):1485-95.
61. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology. 1988;166(1 Pt 1):193-9.
62. Weishaupt D, Zanetti M, Boos N, Hodler J. MR imaging and CT in osteoarthritis of the lumbar facet joints. Skeletal Radiol. 1999;28(4):215-9.
63. Borré DG, Borré GE, Aude F, Palmieri GN. Lumbosacral epidural lipomatosis: MRI grading. Eur Radiol. 2003;13(7):1709-21.
64. Pfirrmann CW, Dora C, Schmid MR, Zanetti M, Hodler J, Boos N. MR image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. Radiology. 2004;230(2):583-8.
65. Team RC. R: A Language and Environment for Statistical Computing. 2020.

66. Ensor J, Martin EC, R.D. R. pmsampsize: Calculates the Minimum Sample Size Required for Developing a Multivariable Prediction Model. R package version 102. 2019.
67. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE, Jr., Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. Stat Med. 2019;38(7):1262-75.
68. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976). 2008;33(1):90-4.
69. Kanters TA, Bouwmans CAM, van der Linden N, Tan SS, Hakkaart-van Roijen L. Update of the Dutch manual for costing studies in health care. PLoS One. 2017;12(11):e0187477.
70. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52.
71. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011;45(3):1 - 67.
72. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344-9.
73. Bahar-Ozdemir Y, Sencan S, Ercalik T, Kokar S, Gunduz OH. The Effect of Pre-Treatment Depression, Anxiety and Somatization Levels on Transforaminal Epidural Steroid Injection: A Prospective Observational Study. Pain Physician. 2020;23(3):E273-e80.
74. McGreevy K, Bottros MM, Raja SN. Preventing Chronic Pain following Acute Pain: Risk Factors, Preventive Strategies, and their Efficacy. Eur J Pain Suppl. 2011;5(2):365-72.
75. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. J Behav Med. 2007;30(1):77-94.
76. Linton SJ. Early identification and intervention in the prevention of musculoskeletal pain. Am J Ind Med. 2002;41(5):433-42.
77. den Boer JJ, Oostendorp RA, Beems T, Munneke M, Evers AW. Continued disability and pain after lumbar disc surgery: the role of cognitive-behavioral factors. Pain. 2006;123(1-2):45-52.
78. Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC. Pathobiology of Modic changes. Eur Spine J. 2016;25(11):3723-34.
79. Crockett MT, Kelly BS, van Baarsel S, Kavanagh EC. Modic Type 1 Vertebral Endplate Changes: Injury, Inflammation, or Infection? AJR Am J Roentgenol. 2017;209(1):167-70.
80. Peterson CK, Pfirrmann CW, Hodler J. Are Modic changes related to outcomes in lumbar disc herniation patients treated with imaging-guided lumbar nerve root blocks? Eur J Radiol. 2014;83(10):1786-92.

7

The effect of a transforaminal epidural injection in patients with lumbar disc herniation is not correlated with the presence of type II Modic changes

E.J.A. Verheijen, N.R.E. van der Vlist, E.C. Bartels,
O.B.H.A.M. van Haagen, C.L.A. Vleggeert-Lankamp

Brain Spine 2025

ABSTRACT

Introduction

Transforaminal epidural steroid injections (TEI) have been suggested to alleviate symptoms in patients with lumbar disc herniation (LDH) through its anti-inflammatory effect. However, treatment effect varies among patients and reliable predictors are lacking. Modic changes (MC) are also associated with inflammatory processes and, therefore, we hypothesize that MC may be correlated with outcome after TEI.

Research question

To investigate the correlation between the presence of MC at the level of LDH and the effect of TEI.

Material and methods

Patients with unilateral lumbar radiculopathy secondary to LDH undergoing TEI were included. MC was graded by two independent assessors. Outcome measures included leg pain, back pain, disability and patient-received recovery at baseline, 30 min, 2 weeks and 6 weeks after treatment. Multivariate analysis was performed for all outcomes and for dichotomized scores using a cutoff of ≥30% improvement. A p-value of ≤0.05 was considered statistically significant.

Results

A total of 88 patients were included of whom 52.3% demonstrated MC. The vast majority was classified as type II (94%). The presence of MC was not correlated with any outcome measure when correcting for age, gender, duration of symptoms and the use of analgesics, nor for dichotomized scores.

Discussion and conclusion

The findings indicate that type II MC is not associated with outcome within six weeks after TEI. Therefore, type II MC cannot be used as a predictor for TEI outcome. Future studies should include longer follow-up and investigate the correlation between the type of MC and the effect of TEI.

INTRODUCTION

Low back pain and unilateral lumbar radiculopathy are common spinal problems associated with disc herniation in the lumbar spine (LDH) (Waldman and Waldman, 2019). Lifetime prevalence rates of up to respectively 80% and 43% have been reported for these conditions (Hoy et al., 2012; Konstantinou and Dunn, 2008) and patients can experience a severe decline of physical functioning as a result. Although symptoms usually spontaneously diminish with time, the period until full recovery can be very debilitating when the pain is insufficiently alleviated with oral analgesics.

In these cases, additional treatment with a Transforaminal Epidural Steroid Injection (TEI) can help to reduce pain symptoms. It is assumed that radicular pain is caused by a combination of mechanical compression, inflammatory reactions and immunological processes affecting the lumbar nerve root (Stafford et al., 2007). Epidural injections with a corticosteroid aim to reduce the inflammation around the nerve root by limiting nerve root oedema, blocking prostaglandin synthesis and altering the conduction of nociceptive C-fibres (McLain et al., 2005; Collighan and Gupta, 2009; Johansson et al., 1990; Kantrowitz et al., 1975). Therefore, this minimally invasive treatment may reduce radicular symptoms and enable the patient to maintain physical functionality, restore quality of life, and, if satisfactorily effective, obviate the need for neurosurgical intervention.

Despite the potential benefits, treatment results with TEI in patients with LDH vary considerably and, therefore, factors that can differentiate between responders and non-responders to TEI would be valuable to achieve personalized treatment strategies. Vertebral end plate changes, or Modic changes (MC), have been suggested to reflect the presence of inflammatory processes and are associated with disc herniation and less favorable outcomes after surgery (Dudli et al., 2017, 2018, 2022; Viswanathan et al., 2020; Heggli et al., 2023; Modic et al., 1988; Albert et al., 2008; Djuric et al., 2019). It is postulated that the absence of MC is associated with less radicular pain and better clinical outcome, whereas the presence of MC aggravates radicular symptoms (Djuric et al., 2019). Hence, we hypothesize that in patients with MC at the level of disc herniation the inflammatory component may have a larger contribution to the radicular symptoms of the patient and, therefore, TEI with anti-inflammatory medication may be more effective. To the best of our knowledge, only two studies have assessed this association before, but demonstrated contradictory results (Lee et al., 2013; Peterson et al., 2014).

Therefore, in this prospective cohort study, we aim to determine the effect of TEI on various clinical outcomes in patients with unilateral radiculopathy secondary to LDH stratified by the presence or absence of MC.

MATERIAL AND METHODS

Patient inclusion

Patients were included that participated in the POTEISS study, an ongoing large prospective cohort study aimed at developing a prediction model for TEI success for patients with lumbar radiculopathy due to LDH or degenerative spinal stenosis. Patients who were referred by the neurologist or neurosurgeon to the outpatient pain clinic in the Spaarne Hospital, Hoofddorp and Haarlem, the Netherlands, for treatment with TEI were eligible for inclusion in this study if clinical findings were in accordance with Magnetic Resonance Imaging (MRI) examination. Patients younger than 18 years, with severe multisegmental spinal degeneration, anatomical abnormalities that may complicate TEI treatment technically (e.g. severe scoliosis), a history of lower back surgery at the same lumbar level and side, previous TEI treatment for their current episode of lumbar radiculopathy, active malignancy or infectious disease, the use of immunosuppressive drugs or systemic corticosteroids in the preceding 3 months, pregnancy, circumstances preventing treatment with TEI (e.g. use of anticoagulants that cannot be temporarily discontinued) or a major language barrier were excluded. Patients were contacted by phone one week before their scheduled TEI appointment and informed about the study. If the patient was willing to participate, they were contacted again two days before treatment for informed consent and to fill in baseline questionnaires.

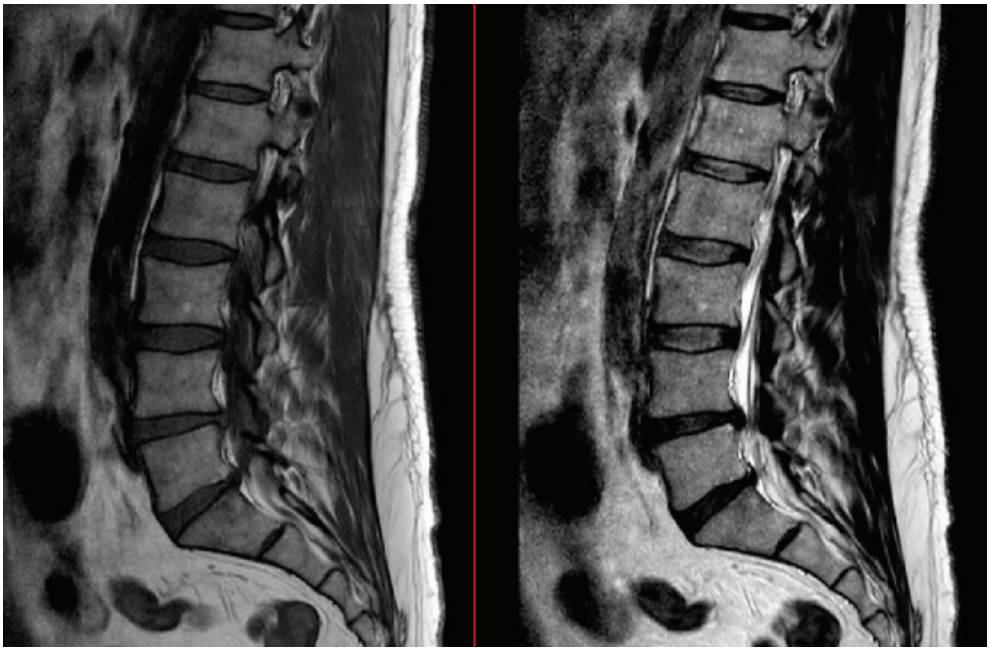
Data collection

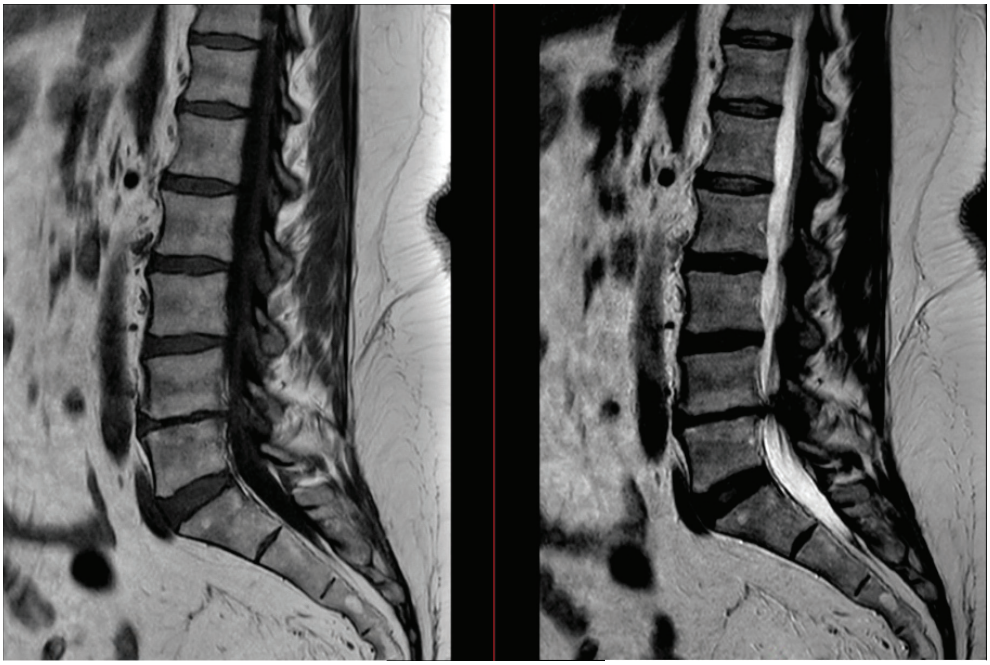
Demographical data were collected at baseline and processed using Castor EDC, a web-based data capture platform (Ciwit B.V., Amsterdam, The Netherlands). Clinical data included the average leg pain score for the past seven days (Numerical Rating Scale (NRS)), average back pain for the past seven days (NRS) and Oswestry Disability Index (ODI) for physical functionality and were gathered at baseline, 2 and 6 weeks after TEI treatment. In addition, for leg and back pain NRS scores at 30 min after the procedure were collected by one of the outpatient pain clinic nurses. Furthermore, patient perceived recovery (GPE) was measured at 2 and 6 weeks follow-up on a 7-point Likert scale.

MC grading

For each patient, the T1-and T2-weighted sagittal MRI scans were assessed by two independent readers (CVL and EV) to determine the level of LDH, presence or absence, type (I, II or III) and severity of MC (mild, moderate or severe) at the same level of the LDH. Grading of MC was performed according to the original classification by Modic et al.: type 1 changes were hypointense on T1-weighted imaging (T1WI) and hyperintense on T2-weighted imaging (T2WI), type 2 changes were hyperintense on T1WI and isointense to hyperintense on T2WI, and type 3 changes were hypointense on both T1WI and T2WI (Modic et al., 1988). If a mixture of MC types was shown, it was categorized according to the most abundant type at the LDH level. The severity of MC was determined by the degree of endplate involvement: mild (<25%), moderate (25–75%) and severe (>75%). In case of discrepancies, consensus was reached through discussion. An example of patient without MC and with MC is shown in Fig. 1.

Fig. 1 Disc herniation at L4-5. Upper panel: Sagittal T1-and T2-weighted MRI images showing normal bone marrow intensity and absence of Modic changes. Lower panel: Sagittal T1-and T2-weighted MRI images showing hyperintense areas around the L4-5 disc, indicating type II Modic changes.





Treatment procedure

Administration of TEI was performed by an anesthesiologist who determined the treatment level based on clinical findings and radiological imaging. The patient lay in a prone position on the table with a support cushion underneath the abdominal region to reduce the natural lordosis. The skin was sterilized with chlorhexidine and locally anaesthetized with 1 mL 1% lidocaine. Under fluoroscopic image guidance the needle was placed via a transforaminal approach in close proximity of the nerve root. Contrast agent was injected to confirm correct positioning of the needle. In accordance with current Dutch anesthesiologic guidelines, injections at L3 and below contained 1,5 ml lidocaine 2% and 40 mg methylprednisolone acetate and injections above L3 contained 1,5 ml lidocaine 1% and 10 mg dexamethasone to prevent vascular occlusion (Dietrich et al., 2015). After the procedure, the patient was monitored for 30 min at the recovery room.

Sample size

The minimal sample size required to assess the correlation between the presence of Modic changes and the effect of TEI was calculated using a two-tailed α of 0.05 and β of 0.2. We aimed to include a sufficient number of patients that would allow to determine whether a correlation coefficient indicating low correlation ($r = 0.3$) differed from zero. Therefore, a total sample size of at least 85 patients was necessary (Hulley et al., 2013).

Data analysis

Statistical analysis was performed using SPSS version 29 (IBM Corp., Armonk, New York, USA). Demographical data were analyzed using descriptive statistics for patients with and without MC separately using appropriate tests depending on the normality of distribution. Linear mixed models were employed to assess the association between the presence of MC (as main effect) and change scores in leg pain, back pain and ODI at follow-up correcting for baseline differences and multiple testing (Holm-Bonferroni correction). Age, gender, duration of symptoms and the use of pain medication at baseline were used as covariates. Additionally, linear mixed models analysis was conducted using change scores expressed in percentages compared to baseline and generalized estimating equations analysis for dichotomized scores for leg pain, back pain, ODI (a decrease of 30% or more compared to baseline was categorized as 'success') and patient perceived recovery (a score of 1 or 2 was considered as 'success'). Loss-to-follow-up was reported for every outcome variable at the various time points and accounted for using statical analysis methods that can handle missing data. A p-value <0.05 was considered to be statistically significant.

RESULTS

A total of 88 patients was included in this study with 52.3% (n = 46) showing MC at the level of LDH. Type II MC was most common with a prevalence of 93.5% (n = 43). At baseline, patients with MC had less leg and back pain, while ODI did not differ. Other demographic variables were not different between groups (Table 1).

Table 1 Baseline characteristics for patients with and without MC at the level of LDH. Continuous variables are presented as mean ± standard deviation; categorical variables are presented in number (percentage).

Variable	Modic changes (n = 46)	No Modic changes (n = 42)	P-value
Age (yrs)	56.9 ± 13.6	50.6 ± 17.2	0.058
Gender (male)	20 (43.5%)	21 (50.0%)	0.540
BMI (kg/m^2)	25.9 ± 3.0	26.0 ± 3.6	0.936
Duration of symptoms (wks)	47.9 ± 70.9	32.0 ± 52.4	0.239
Smoking (yes)	10 (21.7%)	13 (31.0%)	0.326
Alcohol usage			0.254
None	13 (28.3%)	19 (45.2%)	
1-6 units/wk	24 (52.2%)	17 (40.5%)	
>6 units/wk	9 (19.6%)	6 (14.3%)	
Side of leg pain (left)	24 (52.2%)	21 (50.0%)	0.839

Table 1 Baseline characteristics for patients with and without MC at the level of LDH. Continuous variables are presented as mean ± standard deviation; categorical variables are presented in number (percentage). (*continued*)

Variable	Modic changes (n = 46)	No Modic changes (n = 42)	P-value
Use of pain medication (yes)	39 (84.8%)	36 (85.7%)	0.902
Paracetamol	34 (73.9%)	30 (71.4%)	
NSAID	20 (43.5%)	18 (42.9%)	
Morphine	17 (37.0%)	12 (28.6%)	
Physical therapy (yes)	12 (26.1%)	16 (38.1%)	0.227
Level of LDH			0.408
L2-L3	1 (2.2%)	1 (2.4%)	
L3-L4	4 (8.7%)	5 (11.9%)	
L4-L5	21 (45.7%)	25 (59.5%)	
L5-S1	20 (43.5%)	11 (26.2%)	
Type of Modic change			
Type I	2 (4.3%)		
Type II	43 (93.5%)		
Type III	1 (2.2%)		
Severity of Modic change			
Mild	28 (60.9%)		
Moderate	15 (32.6%)		
Severe	3 (6.5%)		
Baseline leg pain	5.6 ± 1.9	6.9 ± 1.3	< 0.001
Baseline back pain	4.0 ± 2.4	5.2 ± 2.2	0.019
Baseline ODI	42.1 ± 18.4	43.0 ± 15.8	0.806

Leg pain, back pain and ODI scores decreased during 6 weeks for both groups (Fig. 2, Fig. 3, Fig. 4). Data were missing for some patients depending on the outcome variable and follow-up time point (Table 2).

Fig. 2 Comparison of the effect of TEI on leg pain between patients with MC and without MC at the level of LDH.

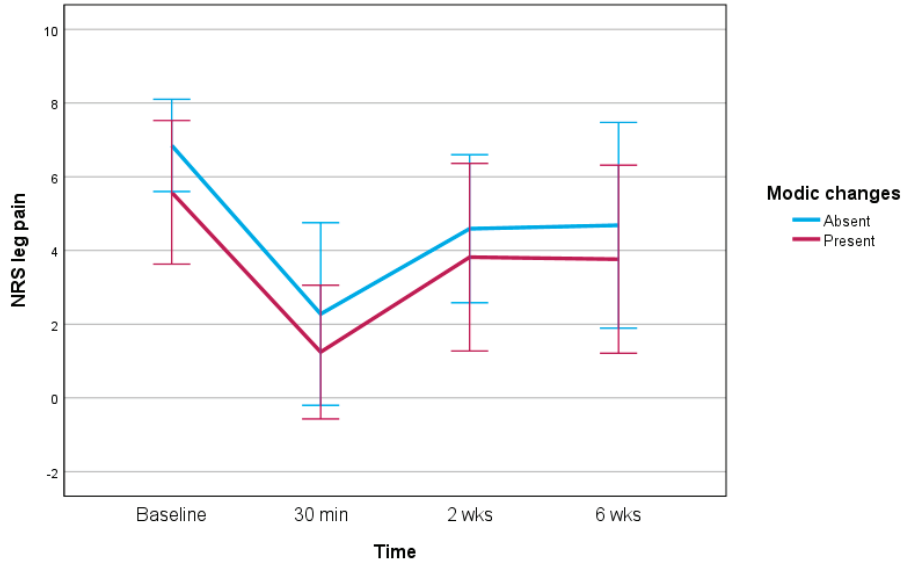


Fig. 3 Comparison of the effect of TEI on back pain between patients with MC and without MC at the level of LDH.

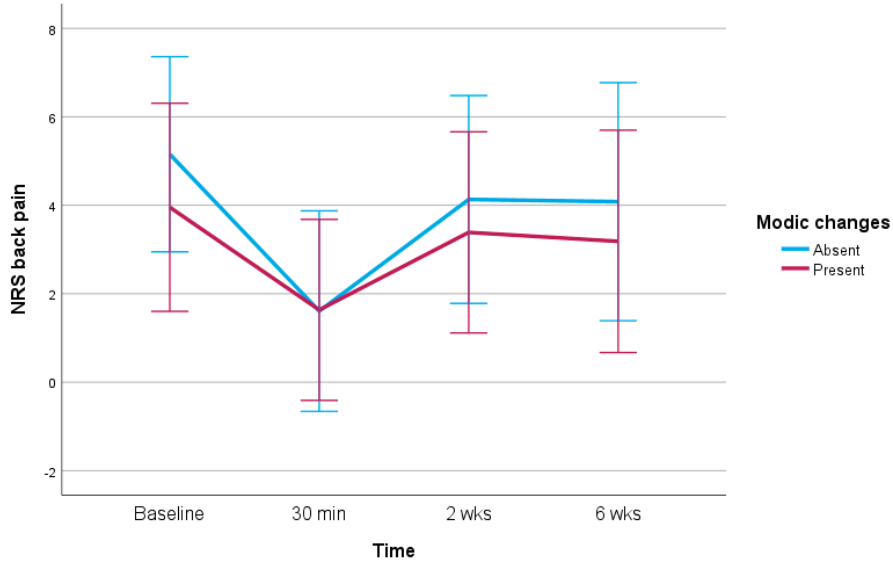


Fig. 4 Comparison of the effect of TEI on functional status between patients with MC and without MC at the level of LDH.

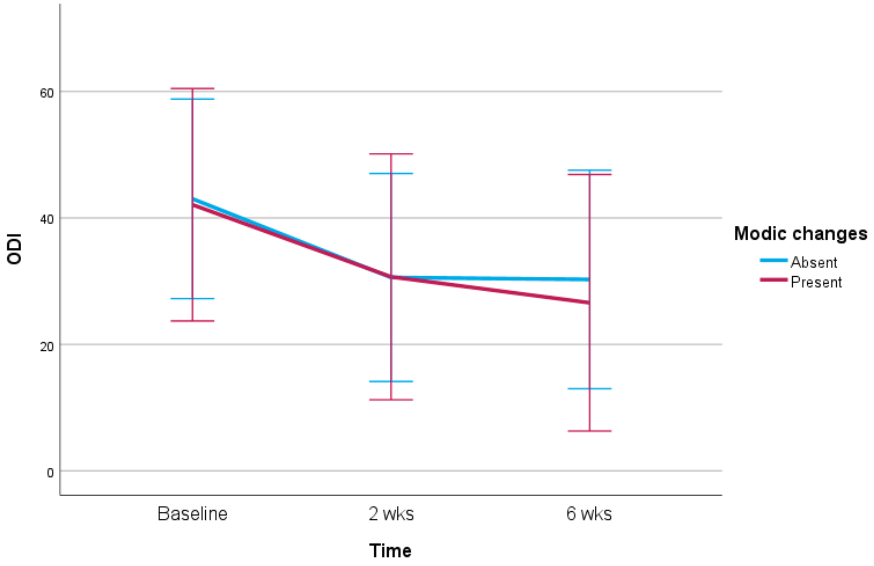


Table 2 Missing data for each outcome variable at all follow-up time points.

FU	NRS leg pain	NRS back pain	ODI	GPE
30 min	19/88 (21.6%)	21/88 (23.9%)	-	-
2 wks	6/88 (6.9%)	7/88 (8.0%)	7/88 (8.0%)	8/88 (9.1%)
6 wks	13/88 (14.8%)	14/88 (15.9%)	14/88 (15.9%)	15/88 (17.0%)

GPE: Global Perceived Effect (recovery); NRS: Numerical Rating Scale; ODI: Oswestry Disability Index.

Changes in leg pain scores compared to baseline were not associated with the presence or absence of MC when controlling for age, gender, duration of symptoms and the use of analgesic medication (estimate of the fixed effect −0.429; standard error 0.315; 95% CI -1.051, 0.194; p-value 0.176). When converting absolute change scores to changes expressed as a percentage of the baseline score, no association with MC was demonstrated either (estimate −3.209; SE 5.185; 95% CI -7.021, 13.43; p-value 0.537). Back pain was neither associated with MC at the level of LDH for absolute change scores (estimate −0.471; SE 0.353; 95% CI -1.168, 0.225; p-value 0.184) nor percentual changes scores (estimate 4.483; SE 12.310; 95% CI -19.872, 28.838; p-value 0.716). ODI representing the patient's functional status was not associated with MC for both absolute and percentual change scores compared to baseline (estimate 0.697; SE 2.236; 95% CI -3.724, 5.118; p-value 0.756 and estimate 5.292; SE 5.608; 95% CI -5.793, 16.377; p-value 0.347).

All continuous outcome variables were dichotomized using a cut-off of a 30% or more decrease compared to baseline to indicate 'success' after TEI treatment. The presence of Modic changes was not associated with the number of patients categorized as 'success' based on leg pain scores (estimate 0.083; SE 0.3022; 95% CI -0.509, 0.676; p-value 0.783), back pain scores (estimate −0.195; SE 0.3247; 95% CI -0.832, 0.441; p-value 0.548), nor scores for functional status (estimate 0.084; SE 0.3671; 95% CI -0.636, 0.052; p-value 0.820). In addition, dichotomized patient perceived recovery was neither associated with the presence or absence of MC (estimate −0.186; SE 0.3936; 95% CI -0.958, 0.585; p-value 0.636) (Table 3).

Table 3 Number of patients (percentages) that demonstrated a decrease of 30% or more compared to baseline for leg pain, back pain and ODI or considered their recovery 'successful' after treatment with TEI.

	Leg pain		Back pain		ODI		GPE	
	MC	**No MC**	**MC**	**No MC**	**MC**	**No MC**	**MC**	**No MC**
30 min	35/40 (87.5%)	22/29 (75.9%)	28/40 (70.0%)	20/27 (74.1%)	-	-	-	-
2 wks	22/43 (51.2%)	20/39 (51.3%)	19/43 (44.2%)	16/38 (42.1%)	19/43 (44.2%)	16/38 (42.1%)	12/43 (27.9%)	12/37 (32.4%)
6 wks	18/37 (48.6%)	21/37 (56.8%)	13/38 (34.2%)	19/36 (52.8%)	21/38 (55.3%)	19/36 (52.8%)	16/38 (42.1%)	16/35 (45.7%)

GPE: Global Perceived Effect (recovery); NRS: Numerical Rating Scale; ODI: Oswestry Disability Index.

DISCUSSION

This study has demonstrated that the presence or absence of Modic changes at the level of disc herniation is not associated with changes in leg pain, back pain, physical functionality nor patient-perceived recovery within six weeks after TEI treatment. Conclusions should be narrowed to type II MC, since 94% of patients with MC had type II MC. In both groups leg pain and back pain decreased over time, and physical functionality increased, with the largest change between baseline and two weeks after treatment, but differences between groups were not statistically significant. These results did not change when considering change scores expressed in percentages or when dichotomized to 'success' or 'non-success'.

We hypothesized that in patients with MC at the level of LDH inflammatory processes may play a larger role and aggravate radicular symptoms causing TEI to be more effective. However, our results demonstrated no correlation between

the effect of TEI and presence of MC. Possibly, MC may not be a marker for nerve root inflammation although it has been associated with disc herniation and less favorable outcomes after disc surgery (Albert et al., 2008; Djuric et al., 2019). More likely, the correlation may have been obscured by the preponderance of MC type II among our study population. Although both type I and type II, the most prevalent types of MC, are associated with inflammatory changes, it is thought that the pathophysiological mechanisms are different. Type II is assumed to have a more chronic character, which fits to the mean duration of symptoms of 32 and 48 weeks in the studied patient groups. It is possible that TEI is less effective in those patients, which might explain the lower percentage of responders among our patients compared to Joswig et al. (48.6–56.8% vs. 66.7%), who reported a mean duration of symptoms of 14.8 weeks (Joswig et al., 2016). Therefore, the effect of TEI may be more pronounced when comparing patients without MC to those with MC type I, as the latter represents a more acute course of symptoms. Furthermore, it is possible that the inflammatory processes related to MC type II may take time to respond to the injection, and these effects may not be noticeable in the early post-injection period, thus requiring longer follow-up.

Our findings agree with the results demonstrated by Lee et al. (2013). Among a set of 149 patients that received up to three transforaminal injections for lumbar radiculopathy caused by LDH, MRI parameters were investigated in a retrospective cohort study. Only patients with the very best or very worst outcome after treatment were included. 13% of patients had MC, 74% of them demonstrated type II MC, and the majority of the patients had symptoms for less than one month. The authors found no correlation between the presence or absence of Modic changes and response after TEI measured by the Visual Analogue Scale (VAS) and a 5-point self-satisfaction scale. However, for their analyses, they did not specify the average time between treatment and outcome assessment. Contrastingly, Peterson et al. did observe a correlation between the presence of Modic changes and pain severity after TEI in their prospective cohort study (Peterson et al., 2014). They included a total of 346 patients; 57% of them had MC and 67% of these patients demonstrated type II MC. They found that patients with MC (type I and II combined) reported significantly lower levels of pain reduction compared to patients without MC at one month follow-up ($p = 0.04$), but not for shorter follow-up nor for dichotomized Patient's Global Impression of Change (PGIC). The percentage of patients that considered themselves improved, 1 month after TEI (40–50%) was comparable to our results (42–46% after 6 weeks). However, their methods were slightly different as patients were included presenting with MC at any level and not necessarily at the level of disc herniation. They hypothesized that there could be a difference in outcomes

between patients with and without MC, which has been associated with low back pain, as patients with MC would have an additional pain generator. Yet, in our opinion, TEI is aimed at relieving radicular symptoms rather than back pain, and the authors did not specify whether the NRS pain scores patients reported concerned leg pain or back pain. Moreover, patients could receive injections for multiple nerve roots and statistical analyses did not correct for any confounders. In addition, very few baseline characteristics were presented for the patient groups with and without MC as reliable information on socio-demographic factors was not available rendering a multivariate analysis in their study infeasible. Finally, 65 patients in this cohort demonstrated type I MC, which may have contributed to a different outcome.

The results from our study are strengthened by the strict in- and exclusion criteria applied during inclusion of patients. Only patients with unilateral radiculopathy and clinical symptoms that were concordant with radiological findings on MRI, were eligible to participate. Furthermore, treatment was limited to a single fluoroscopy-guided injection and no repeat injections were performed during follow-up. This allowed for a more appropriate evaluation of TEI effect compared to studies that included patients with bilateral symptoms, without radiological confirmation of LDH, offered bilateral or multi-level injections at baseline or repeat injections during follow-up. Moreover, the prospective nature of this study allowed for the use of validated outcome instruments to systematically collect data on the effect of TEI treatment. Finally, the dichotomization of NRS scores for leg pain was performed using a cut-off that has been defined by an international consensus group of experts to obtain success data (Ostelo et al., 2008). The use of dichotomized data can provide other insights on the effect of TEI, since it is possible that the true effect is obscured using mean data from a continuous variable (Ghahreman et al., 2010).

The current study has a few limitations. First, the total number of patients included was small since this was a subset of a larger cohort study. For some patients outcome data was missing which reduced the number of complete cases for our analyses. Furthermore, the vast majority of patients with Modic changes demonstrated type II which precluded a sub-analysis of the type of MC on the effect of TEI.

Future studies should include a larger sample size to obtain a higher statistical power and to investigate the correlation between MC type and the effect of TEI. In addition, more patients with a shorter duration of symptoms should be included to determine whether MC type I is more prevalent in this group and

if the outcome after TEI differs compared to those with MC type II and those without MC.

The findings from our study indicate that there is no correlation between the presence or absence of type II Modic changes at the same level of LDH and outcome after TEI on the short term. Therefore, clinicians should not consider the presence of type II MC to be a contra-indication for treatment with TEI.

REFERENCES

[1] Waldman SD. 84 - Spinal Stenosis. In: Waldman SD, editor. Atlas of Common Pain Syndromes (Fourth Edition). Philadelphia: Elsevier; 2019. p. 324-7.

[2] Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. Arthritis Rheum. 2012;64(6):2028-37.

[3] Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. Spine (Phila Pa 1976). 2008;33(22):2464-72.

[4] Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth. 2007;99(4):461-73.

[5] McLain RF, Kapural L, Mekhail NA. Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J. 2005;5(2):191-201.

[6] Collighan N, Gupta S. Epidural steroids. Continuing Education in Anaesthesia Critical Care & Pain. 2009;10(1):1-5.

[7] Johansson A, Hao J, Sjölund B. Local corticosteroid application blocks transmission in normal nociceptive C-fibres. Acta Anaesthesiol Scand. 1990;34(5):335-8.

[8] Kantrowitz F, Robinson DR, McGuire MB, Levine L. Corticosteroids inhibit prostaglandin production by rheumatiod synovia. Nature. 1975;258(5537):737-9.

[9] Dudli S, Miller S, Demir-Deviren S, Lotz JC. Inflammatory response of disc cells against Propionibacterium acnes depends on the presence of lumbar Modic changes. European Spine Journal. 2018;27(5):1013-20.

[10] Viswanathan VK, Shetty AP, Rajasekaran S. Modic changes - An evidence-based, narrative review on its patho-physiology, clinical significance and role in chronic low back pain. J Clin Orthop Trauma. 2020;11(5):761-9.

[11] Heggli I, Laux CJ, Mengis T, Karol A, Cornaz F, Herger N, et al. Modic type 2 changes are fibroinflammatory changes with complement system involvement adjacent to degenerated vertebral endplates. JOR Spine. 2023;6(1):e1237.

[12] Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology. 1988;166(1 Pt 1):193-9.

[13] Dudli S, Karol A, Giudici L, Heggli I, Laux CJ, Spirig JM, et al. CD90-positive stromal cells associate with inflammatory and fibrotic changes in modic changes. Osteoarthr Cartil Open. 2022;4(3):100287.

[14] Dudli S, Sing DC, Hu SS, Berven SH, Burch S, Deviren V, et al. ISSLS PRIZE IN BASIC SCIENCE 2017: Intervertebral disc/bone marrow cross-talk with Modic changes. European Spine Journal. 2017;26(5):1362-73.

[15] Albert HB, Kjaer P, Jensen TS, Sorensen JS, Bendix T, Manniche C. Modic changes, possible causes and relation to low back pain. Med Hypotheses. 2008;70(2):361-8.

[16] Djuric N, Yang X, Ostelo R, van Duinen SG, Lycklama À Nijeholt GJ, van der Kallen BFW, et al. Disc inflammation and Modic changes show an interaction effect on recovery after surgery for lumbar disc herniation. Eur Spine J. 2019;28(11):2579-87.

[17] Lee JW, Choi SW, Park SH, Lee GY, Kang HS. MR-based outcome predictors of lumbar transforaminal epidural steroid injection for lumbar radiculopathy caused by herniated intervertebral disc. Eur Radiol. 2013;23(1):205-11.

[18] Peterson CK, Pfirrmann CW, Hodler J. Are Modic changes related to outcomes in lumbar disc herniation patients treated with imaging-guided lumbar nerve root blocks? Eur J Radiol. 2014;83(10):1786-92.

[19] Dietrich TJ, Sutter R, Froehlich JM, Pfirrmann CW. Particulate versus non-particulate steroids for lumbar transforaminal or interlaminar epidural steroid injections: an update. Skeletal Radiol. 2015;44(2):149-55.
[20] Hulley SB CS, Browner WS, Grady D, Newman TB. Designing clinical research : an epidemiologic approach. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.
[21] Joswig H, Neff A, Ruppert C, Hildebrandt G, Stienen MN. The Value of Short-Term Pain Relief in Predicting the One-Month Outcome of Lumbar Transforaminal Epidural Steroid Injections. World Neurosurg. 2016;96:323-33.
[22] Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976). 2008;33(1):90-4.
[23] Ghahreman A, Ferch R, Bogduk N. The efficacy of transforaminal injection of steroids for the treatment of lumbar radicular pain. Pain Med. 2010;11(8):1149-68.

Part II

Improving diagnosis and prognosis in lumbar spinal stenosis

8

General introduction and outline

Lumbar spinal stenosis (LSS) is a degenerative condition characterized by the narrowing of the spinal or nerve root canal, resulting in compression of neural structures. Stenosis is generally classified into three subtypes based on its anatomical location: (1) central canal stenosis (CCS), caused by disc bulging and ligamentum flavum hypertrophy; (2) lateral recess stenosis (LRS), resulting from disc bulging and facet joint arthropathy (FA); and (3) foraminal stenosis (FS), attributed to degenerative changes within the neuroforamen. CCS typically results in neurogenic claudication, characterized by bilateral radiating leg pain that worsens with walking or lumbar extension and improves with sitting or lumbar flexion. In contrast, LRS and FS more frequently present with neurogenic claudication symptoms with a radicular pattern. Often accompanied by back pain, LSS can significantly impair mobility and quality of life [1, 2]. Given its degenerative nature, LSS is most common in the elderly. As such, its prevalence is increasing due to our aging population, and, hence, it is one of the most frequent indications for spinal surgery in older adults [3, 4]. However, despite advances in imaging and surgical techniques, clinical decision-making regarding LSS remains challenging, and postoperative patient satisfaction varies widely, ranging from 27% to 80% [5-7].

One of the significant barriers to clinical decision-making for LSS is the ambiguity in radiological assessment. Magnetic resonance imaging (MRI) is the gold standard for diagnosing LSS, offering high-resolution visualization of soft tissues and neural compression [8]. In addition to MRI, computed tomography (CT) and plain radiography are utilized to evaluate bony structures. Standardized assessment of the severity of LSS on diagnostic imaging is essential for unequivocal communication between radiologists, neurologists and neurosurgeons, and for guiding appropriate treatment strategies. To achieve this, grading systems must be capable of producing high inter-reader agreement, clinically applicable—avoiding excessive complexity while maintaining sufficient detail for meaningful differentiation—and correlate with clinical symptoms [9].

The first widely adopted classification system for LSS, introduced by Lurie et al., categorized CCS as mild, moderate, or severe [10]. Subsequently, more detailed grading systems were developed by Guen et al. and Schizas et al., but these also focused exclusively on the central canal [11, 12]. Although moderate to almost perfect inter-reader agreement has been reported, external validation studies have demonstrated substantial variability in interobserver reliability. Moreover, these grading systems have shown inconsistent correlations with clinical symptoms and do not account for stenosis of the lateral recess (LRS) or neuroforamen (FS). Conversely, grading systems developed specifically for LRS and FS share similar limitations and do not incorporate assessment of CCS. To

address these limitations, Miskin et al. introduced a novel, more comprehensive grading system in 2021 [13]. This classification extends assessment beyond CCS to include LRS, FS, and FA for a more complete evaluation of LSS. While this system demonstrated fair to substantial inter-reader agreement, it has yet to undergo external validation or be correlated with baseline clinical symptoms and postoperative outcomes.

Although some grading scales have demonstrated adequate inter-reader agreement, human variability in interpretation can never be completely eliminated. Additionally, certain grading systems are laborious, limiting their utility in the fast-paced clinical environment. Automating the grading process could address both challenges. Artificial intelligence (AI), particularly machine learning (ML) and its subtype deep learning (DL), has shown significant promise in medical imaging applications, including automated image segmentation and classification [14, 15]. In segmentation, each pixel is assigned to a class based on extracted attributes, which are then used as inputs for classification models that predict LSS severity. AI-driven models can rapidly analyze MRI scans with high accuracy and consistency, reducing observer variability and potentially enhancing diagnostic precision. While conventional ML approaches require manual selection of relevant input features for segmentation and classification, DL models can autonomously learn relevant imaging features, obviating the need for a separate segmentation step and making them particularly well-suited for complex image analysis tasks.

Beyond diagnosis, radiological imaging may also aid in predicting patient outcomes following surgical treatment for LSS. Previous research has shown that the grading system by Schizas et al. lacks a strong correlation with postoperative outcomes [11, 16, 17]. While some radiological, demographic, and clinical parameters have been associated with surgical outcomes, they are not considered established outcome predictors [16, 18, 19]. Hence, a reliable grading system that is also clinically relevant is still lacking.

AIMS AND OUTLINE OF THIS THESIS:

1. A comprehensive MRI-based classification of LSS severity requires assessment of stenosis not only in the central canal but also in the lateral recess and neuroforamen. The grading system introduced by Miskin et al. incorporates all these anatomical regions and has demonstrated fair to substantial inter-reader agreement. However, further validation is warranted for broader acceptance. Therefore, the first aim of this study is to evaluate inter-reader agreement of this grading system using an independent dataset, providing insights into its validity, and to determine its correlations with baseline symptoms and postoperative outcomes, thereby assessing its clinical relevance and applicability (Chapter 9).
2. Recent advancements in artificial intelligence (AI) have led to the development of sophisticated machine learning algorithms with potential applications in medical imaging. AI-based approaches could improve LSS diagnostics by improving efficiency, reducing observer variability, and lowering healthcare costs. To explore this potential, the second aim of this study is to conduct a systematic review of studies utilizing conventional machine learning (ML) and deep learning (DL) models for the segmentation and classification of LSS, with a focus on evaluating their diagnostic accuracy and performance (Chapter 10).

REFERENCES

[1] Katz JN, Harris MB. Clinical practice. Lumbar spinal stenosis. N Engl J Med. 2008;358(8):818–25.

[2] Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and Management of Lumbar Spinal Stenosis: A Review. Jama. 2022;327(17):1688–99.

[3] Jensen RK, Jensen TS, Koes B, Hartvigsen J. Prevalence of lumbar spinal stenosis in general and clinical populations: a systematic review and meta-analysis. Eur Spine J. 2020;29(9):2143–63.

[4] Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. Jama. 2010;303(13):1259–65.

[5] Weinstein JN, Tosteson TD, Lurie JD, Tosteson A, Blood E, Herkowitz H, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis four-year results of the Spine Patient Outcomes Research Trial. Spine (Phila Pa 1976). 2010;35(14):1329–38.

[6] Haddadi K, Ganjeh Qazvini HR. Outcome after Surgery of Lumbar Spinal Stenosis: A Randomized Comparison of Bilateral Laminotomy, Trumpet Laminectomy, and Conventional Laminectomy. Front Surg. 2016;3:19.

[7] Neukamp M, Perler G, Pigott T, Munting E, Aebi M, Röder C. Spine Tango annual report 2012. Eur Spine J. 2013;22 Suppl 5(Suppl 5):767–86.

[8] Andreisek G, Hodler J, Steurer J. Uncertainties in the diagnosis of lumbar spinal stenosis. Radiology. 2011;261(3):681–4.

[9] Schroeder GD, Kurd MF, Vaccaro AR. Lumbar Spinal Stenosis: How Is It Classified? J Am Acad Orthop Surg. 2016;24(12):843–52.

[10] Lurie JD, Tosteson AN, Tosteson TD, Carragee E, Carrino JA, Kaiser J, et al. Reliability of readings of magnetic resonance imaging features of lumbar spinal stenosis. Spine (Phila Pa 1976). 2008;33(14):1605–10.

[11] Schizas C, Theumann N, Burn A, Tansey R, Wardlaw D, Smith FW, et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphology of the dural sac on magnetic resonance images. Spine (Phila Pa 1976). 2010;35(21):1919–24.

[12] Guen YL, Joon WL, Hee SC, Kyoung-Jin O, Heung SK. A new grading system of lumbar central canal stenosis on MRI: an easy and reliable method. Skeletal Radiology. 2011;40(8):1033–9.

[13] Miskin N, Isaac Z, Lu Y, Makhni MC, Sarno DL, Smith TR, et al. Simplified Universal Grading of Lumbar Spine MRI Degenerative Findings: Inter-Reader Agreement of Non-Radiologist Spine Experts. Pain Med. 2021;22(7):1485–95.

[14] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88.

[15] Wang H, Zhou Z, Li Y, Chen Z, Lu P, Wang W, et al. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from (18)F-FDG PET/CT images. EJNMMI Res. 2017;7(1):11.

[16] Moojen WA, Schenck CD, Lycklama À Nijeholt GJ, Jacobs WCH, Van der Kallen BF, Arts MP, et al. Preoperative MRI in Patients With Intermittent Neurogenic Claudication: Relevance for Diagnosis and Prognosis. Spine (Phila Pa 1976). 2018;43(5):348–55.

[17] Minetama M, Kawakami M, Teraguchi M, Matsuo S, Enyo Y, Nakagawa M, et al. MRI grading of spinal stenosis is not associated with the severity of low back pain in patients with lumbar spinal stenosis. BMC Musculoskelet Disord. 2022;23(1):857.
[18] Aalto TJ, Malmivaara A, Kovacs F, Herno A, Alen M, Salmi L, et al. Preoperative predictors for postoperative clinical outcome in lumbar spinal stenosis: systematic review. Spine (Phila Pa 1976). 2006;31(18):E648–63.
[19] Pearson A, Lurie J, Tosteson T, Zhao W, Abdu W, Weinstein JN. Who should have surgery for spinal stenosis? Treatment effect predictors in SPORT. Spine (Phila Pa 1976). 2012;37(21):1791–802.

9

External validation of a novel comprehensive grading system for lumbar spinal stenosis

E.J.A. Verheijen, R. Lindberg, W.A. Moojen, G.J. Lycklama à Nijeholt,
B.F.W. van der Kallen, C.L.A. Vleggeert-Lankamp

Submitted

ABSTRACT

Purpose

A standardized grading system for lumbar spinal stenosis (LSS) is lacking. This study evaluated inter-reader agreement and clinical utility of the comprehensive grading system proposed by Miskin et al., assessing central canal stenosis (CCS), foraminal stenosis (FS), lateral recess stenosis (LRS), and facet arthropathy (FA).

Methods

Preoperative MRI data from 155 patients with neurogenic claudication due to LSS were retrospectively analyzed. Two neuroradiologists and one spine surgeon independently graded all lumbar levels using the four-item system. Inter-reader agreement was assessed using weighted Cohen's kappa, exact agreement, and agreement within one grade. Associations between radiological grades and baseline symptoms and postoperative outcomes at 8, 26 and 52 weeks were analyzed using bivariate analyses and multivariable linear regression.

Results

Substantial inter-reader agreement was observed for CCS (kappa = 0.751) and LRS (kappa = 0.619), with exact agreement of 61.2% and 68.7%, respectively. In contrast, agreement was slight for FS ($\kappa = 0.214$) and FA ($\kappa = 0.201$). Agreement was highest for extreme grades across all components. At L4-5 and L5-S1, agreement for CCS was comparable to Miskin et al., whereas agreement for LRS was higher. Higher baseline CCS was independently associated with greater postoperative improvement in pain and physical disability, while correlations for the combined CCS+LRS score were weaker.

Conclusion

The Miskin et al. grading system demonstrates substantial reliability for CCS and LRS, but clinical relevance only for CCS. Limited reliability for FS and FA suggests these components require refinement or substitution. Comparative evaluation with alternative grading systems may further enhance clinical applicability.

INTRODUCTION

Over the past few decades, magnetic resonance imaging (MRI) has become the preferred imaging modality for diagnosing patients suspected of having lumbar spinal stenosis (LSS)[1]. Several grading systems have been proposed to assess the degree of LSS on MRI, typically evaluating the central canal zone, lateral recess, or neuroforamen – regions considered clinically relevant [2-12]. These systems aim to standardize LSS assessment and improve intra- and interdisciplinary communication. In addition, standardized grading is more likely to correlate with clinical symptoms and therapeutic outcomes, thereby aiding clinical decision-making.

Lurie et al. were the first to introduce a grading scale for LSS, comprising four grades (none, mild, moderate or severe) based on compromise of the central zone, lateral recess or neuroforamen, expressed in thirds [2]. Since then, more complex grading tools have been developed. The Schizas classification, which remains one of the most widely used grading systems for LSS, consists of four grades ranging from A (no or minor stenosis) to D (extreme stenosis), with grade A further subdivided into four subgrades to assess central canal stenosis (CCS)[3]. Guen et al. subsequently proposed a four-grade scale ranging from no to severe stenosis [4]. Additionally, Weber et al. and Yuan et al. introduced a modified Schizas scale and a novel system, respectively [5, 6]. Further studies proposed grading systems for lateral recess stenosis (LRS), foraminal stenosis (FS), and facet arthropathy (FA) [7-12]. However, most systems are subject to considerable inter-reader variability and lack correlation with clinical parameters, leading to discrepancies in interdisciplinary communication regarding stenosis severity and limiting widespread clinical adoption.

In 2021, Miskin et al. introduced a novel grading system evaluating four key components of lumbar degeneration: central canal stenosis, foraminal stenosis, lateral recess stenosis and facet arthropathy [13]. This system extends the Schizas classification by incorporating additional regions frequently involved in LSS and, alongside the system proposed by Lurie et al., provides a multifaceted assessment. In the original study, three observers independently assessed MRI scans from 50 patients at the L4-5 and L5-S1 levels, demonstrating fair to substantial inter-reader agreement (Cohen's kappa 0.323 – 0.702). This suggests potential value as a comprehensive grading tool applicable across interdisciplinary clinical settings.

However, external validation is necessary for broader acceptance. The aim of this study is to evaluate inter-reader variability, for L4-5 and L5-S1 combined as

well as for all lumbar levels, and to assess correlations with clinical parameters in order to determine the validity and applicability of this grading system.

MATERIAL AND METHODS

Patient population

Patients were included from the FELIX (Foraminal Enlargement Lumbar Interspinous distraXion) trial, a multicenter randomized study comparing interspinous process implants with conventional decompression surgery for intermittent neurogenic claudication due to LSS [14]. The trial demonstrated similar postoperative outcomes between treatment groups. Patients were aged 40-85 years, had symptoms >3 months unresponsive to conservative therapy, underwent surgery at one or two levels and had a preoperative MRI examination. MRI protocols included T1- and T2-weighted sagittal and axial sequences, obtained with 1.5T or 3.0T scanners, slice thickness 3.0-6.0 mm and interslice gaps of 0.3-5.2 mm.

MRI evaluation and grading

Two experienced neuroradiologists (GLN, BvdK) and one experienced spine-dedicated neurosurgeon (CVL) graded all lumbar levels from L1-2 to L5-S1 using Miskin et al.'s system for CCS, FS, LRS and FA [13]. CCS, FS and FA comprised six possible grades, whereas LRS comprised three possible grades. Both FS and LRS were graded separately for the left and right sides. Reviewers were blinded to operated levels and clinical data, and received a presentation with representative examples beforehand to ensure uniform understanding of the grading scale.

Clinical outcome parameters

The Visual Analogue Scale (VAS) for leg and back pain (0-100 scale)[15], and the modified Roland-Morris Disability Questionnaire (mRMDQ) (0-23 scale) [16] were used as outcome parameters at baseline and at 8, 26 and 52 weeks. Change scores were calculated for each follow-up time point.

Inter-reader agreement analysis

Inter-reader agreement was initially assessed for L4-5 and L5-S1 to match the analysis by Miskin et al. Proportions were calculated for all possible grades. Inter-reader agreement was determined using weighted Cohen's kappa with 95% confidence intervals (CIs) and categorized according to Landis and Koch [17]. Exact and within-one-grade agreement between reviewer pairs were calculated. Agreement across stenosis severity categories was analysed using median

reviewer grades to define ordered severity groups. Analyses were repeated for all lumbar levels combined and for individual levels separately. Missing gradings were omitted pairwise.

Correlation with clinical outcomes

Due to low inter-reader agreement, FA and FS were excluded from correlation analyses with clinical outcomes. Analyses therefore focused on CCS and LRS. Median grades were used for operated levels; in case of two surgical levels, the higher median was selected. In addition to the individual grading items, a combined CCS+LRS score was established by recategorizing CCS into 3 grades and summing with LRS, to assess whether a composite measure of stenosis severity demonstrated stronger correlations with clinical outcomes. Correlations were assessed through bivariate analyses with Spearman's rank correlation coefficient and multivariable linear regression models adjusted for age, gender, body mass index (BMI), duration of symptoms, Hospital Anxiety and Depression Scale (HADS) score and treatment group. A p value ≤ 0.05 was considered statistically significant.

RESULTS

A total of 159 patients were enrolled in the FELIX trial and underwent surgery. Preoperative MRI data were available for 155 patients. Baseline characteristics are presented in Table 1.

Table 1 Baseline characteristics

	N	Mean ± std
Age (yrs)	155	66.0 ± 9.1
Gender (male)	155	52.9% (82)
BMI (kg/m²)	155	27.7 ± 4.6
Smoking	153	
None		74.5% (114)
Sporadically		3.9% (6)
Frequently		21.6% (33)
Duration of symptoms (mos)	154	12.0 (6.0 – 24.0)*
VAS leg pain	152	56.1 ± 23.6
VAS back pain	152	49.4 ± 26.3
mRMDQ	142	14.0 ± 4.8
Randomization group	155	
Standard decompression		51.0% (79)
Interspinous device		49.0% (76)

Table 1 Baseline characteristics (*continued*)

	N	Mean ± std
Surgery at second level (yes)	155	21.3% (33)
Primary surgical level	155	
L2-3		4.5% (7)
L3-4		32.9% (51)
L4-5		60.6% (94)
L5-6		0.6% (1)
L4-S1		1.3% (2)
Secondary surgical level	33	
L3-4		12.1% (4)
L4-5		87.9% (29)

BMI: body mass index; mRMDQ: modified Roland-Morris Disability Questionnaire; VAS: visual analogue scale. *Median (interquartile range)

Inter-reader variability analyses for L4-5 and L5-S1

For CCS, 739 readings were available (L4-5: 432; L5-S1: 307). The majority was graded as normal, followed by severe and moderate-severe stenosis (Figure 1). Inter-reader greement was substantial (kappa 0.764; 95% CI 0.733 – 0.795) (Figure 2), with exact agreement in 61.3% (340/555) and within-one-grade agreement in 88.6% (492/555) (Table 2).

Of the 1475 readings for FS (L4-5: 864; L5-S1: 611), most were graded as normal or mild, without major differences between the left and right sides. Inter-reader agreement was slight (kappa 0.204; 95% CI 0.169 – 0.238), with exact agreement in 39.5% (438/1108) and within-one-grade agreement in 64.4% (714/1108). Agreement did not differ significantly between the left and right sides.

For LRS, 1477 readings were available (L4-5: 865; L5-S1: 612), of which most were graded as normal or nerve root compression, again without differences between the left and right sides. Agreement was substantial (kappa 0.610; 95% CI, 0.572 – 0.648), with exact agreement in 69.2% (766/1107) and within-one-grade agreement in 95.8% (1061/1107). Differences between the left and right sides were minimal.

For FA, 733 readings were available (L4-5: 431; L5-S1: 302), which were relatively evenly distributed across grading categories. Agreement was fair (kappa 0.251; 95% CI 0.207 – 0.295), with exact agreement in 25.2% (138/548) and within-one-grade agreement in 60.8% (333/548). Sensitivity analysis evaluating whether collapsing FA grades into three categories improved inter-reader agreement demonstrated a comparable kappa value (0.255; 95% CI 0.203-0.308).

Across CCS, FS, LRS and FA, agreement was highest in the extreme severity categories (Figures 3 & 4).

Inter-reader variability analyses for all lumbar levels

Agreement across all lumbar levels was comparable to that observed at the L4-5 and L5-S1 levels (Table 3). For FS exact agreement was slightly higher when all levels were included, without meaningful effect on kappa values. Collapsing FA grades modestly increased agreement (kappa 0.337; 95% CI 0.297-0.377).

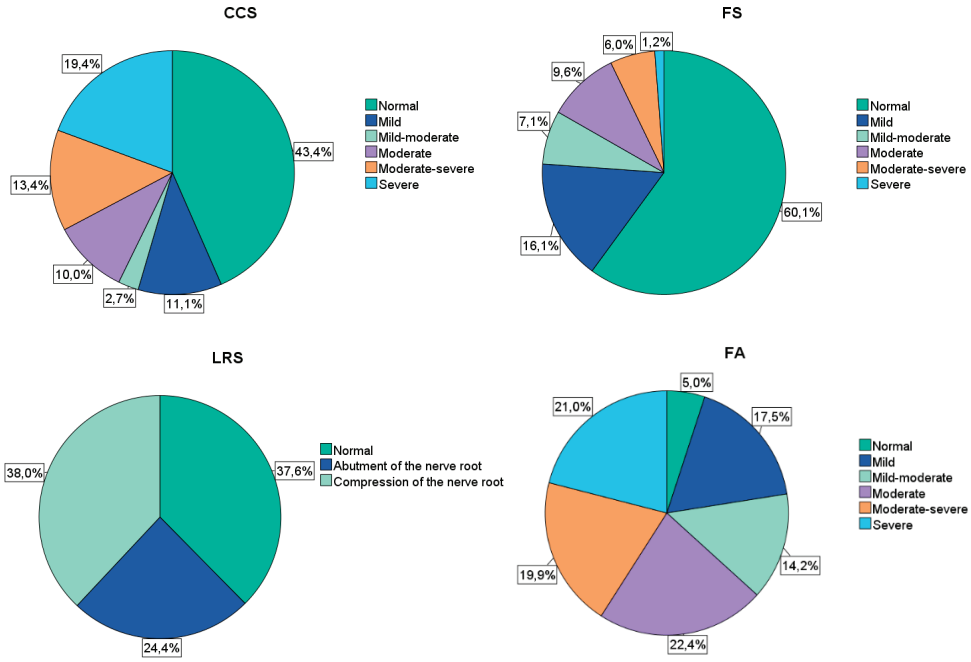


Fig. 1 Distribution of the gradings for CCS, FS, LRS and FA for L4-5 and L5-S1 levels

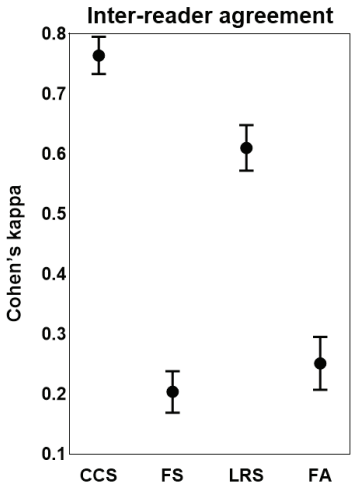


Fig. 2 Box plot depicting the average Cohen's kappa and 95% confidence interval for each item of the scale at L4-5 and L5-S1

Table 2 Inter-reader agreement for CCS, FS, LRS and FA at L4-5 and L5-S1

	Our results			Miskin et al.		
	Kappa (range)	**Exact agreement**	**Agreement within one grade**	**Kappa (95% CI)**	**Exact agreement**	**Agreement within one grade**
CCS	0.764 (0.733-0.795)	61.3%	88.6%	0.702 (0.641 -0.764)	56%	91%
FS	0.204 (0.169-0.238)	39.5%	64.4%	0.544 (0.486 -0.602)	41.5%	78.5%
LRS	0.610 (0.572-0.648)	69.2%	95.8%	0.323 (0.255 -0.392)	34%	90.5%
FA	0.251 (0.207-0.295)	25.2%	60.8%	0.557 (0.495 -0.620)	33%	76%

CCS: central canal stenosis; FA: facet arthropathy; FS: foraminal stenosis; LRS: lateral recess stenosis

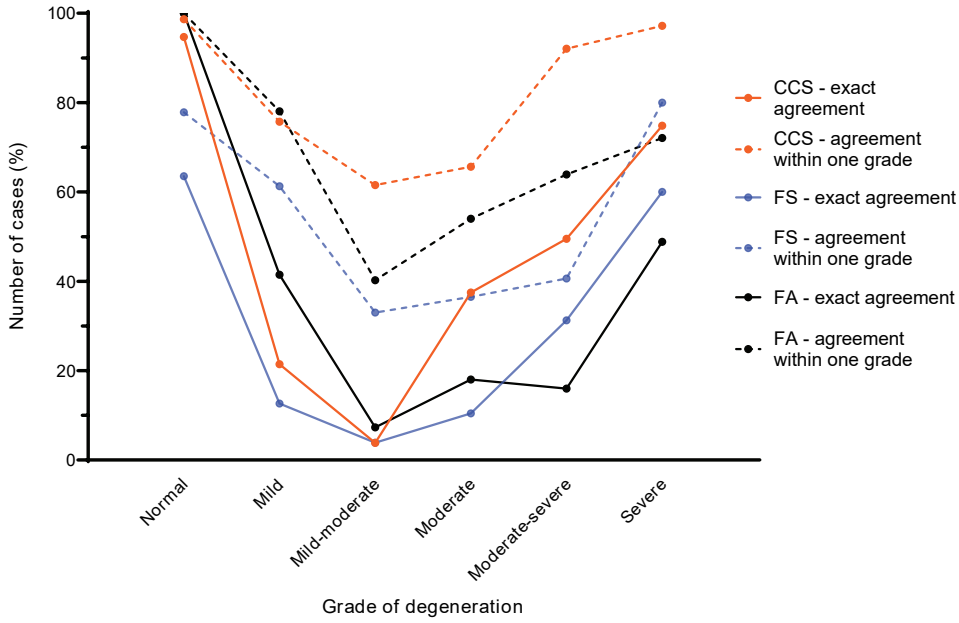


Fig. 3 Distribution of inter-reader agreement by severity of degeneration for CCS, FS and FA.

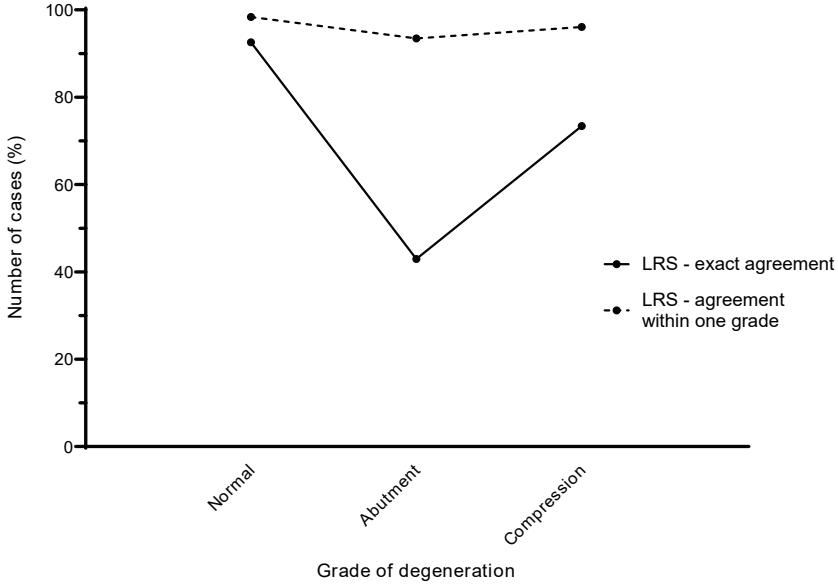


Fig. 4 Distribution of inter-reader agreement by severity of degeneration for LRS.

Table 3 Inter-reader agreement for CCS, FS, LRS and FA for the whole lumbar spine and individual levels

	Whole lumbar spine			L1-2	L2-3	L3-4	L4-5	L5-S1
	Kappa (range)	Exact agreement	Agreement within one grade	Kappa (range)				
CCS	0.751 (0.727-0.776)	61.2% (636/1039)	88.3% (917/1039)	0.524 (0.293-0.754)	0.608 (0.500-0.716)	0.685 (0.629-0.741)	0.639 (0.593-0.686)	0.293 (0.095-0.490)
FS	0.214 (0.187-0.242)	49.0% (1016/2073)	71.4% (1481/2073)	0.237 (-0.033-0.506)	0.071 (-0.068-0.210)	0.131 (0.084-0.177)	0.150 (0.116-0.185)	0.380 (0.279-0.481)
LRS	0.619 (0.592-0.646)	68.7% (1421/2067)	95.4% (1972/2067)	0.585 (0.393-0.777)	0.499 (0.411-0.588)	0.498 (0.438-0.558)	0.371 (0.314-0.429)	0.408 (0.294-0.521)
FA	0.201 (0.102-0.300)	23.6% (241/1020)	62.3% (635/1020)	0.158 (-0.012-0.327)	0.201 (0.102-0.300)	0.243 (0.187-0.299)	0.205 (0.158-0.251)	0.103 (0.014-0.191)

CCS: central canal stenosis; FA: facet arthropathy; FS: foraminal stenosis; LRS: lateral recess stenosis

Correlations between stenosis severity and clinical outcomes

Bivariate analyses for CCS and LRS demonstrated no significant correlations with clinical outcomes at baseline (Table 4). At 8 weeks, weak correlations were observed only between stenosis severity and physical disability (mRMDQ). At 26 and 52 weeks, higher CCS scores were correlated with greater reductions in leg pain, back pain and physical disability, whereas correlations for LRS were weaker and less consistent. The combined CCS+LRS score demonstrated similar, but generally weaker, correlations compared with CCS alone.

Multivariable linear regression analyses demonstrated no significant baseline associations between stenosis severity and clinical parameters (Table 5). However, female sex was independently associated with higher pain scores (VAS leg pain: B = 8.5-8.9, p = 0.024-0.031; VAS back pain: B = 10.2-10.4, p = 0.019-0.021) and greater physical disability (B = 2.8, p<0.001) at baseline. At 8 weeks, higher stenosis severity was associated with greater improvement in physical disability across all grading scores ($p \leq 0.006$). Trend-level correlations were observed between higher CCS and CCS+LRS scores and lower leg pain. At 26 weeks, improvement in physical disability correlated with all three scores ($p \leq 0.015$). Reduction in leg pain correlated only with CCS severity, whereas back pain was associated with CCS and CCS+LRS scores. At 52 weeks, CCS demonstrated the strongest and most consistent associations with all clinical outcomes ($p \leq 0.017$). LRS severity correlated with greater improvement in back pain, showed

no association with leg pain, and demonstrated only a near-significant association with physical disability (p = 0.067). The combined CCS+LRS score was significantly correlated with leg and back pain, whereas the association with physical disability was near-significant (p = 0.077). Treatment group affected outcomes in only a limited number of analyses.

Clinical effect size

Leg pain scores decreased by 5.6 and 7.9 points per one-category increase in CCS severity at 26 and 52 weeks, respectively. For the CCS+LRS score, leg pain decreased by 9.7 points per category at 52 weeks. Back pain scores decreased by 5.8 and 6.5 points per CCS category at 26 and 52 weeks, respectively, compared with decreases of 6.3 and 8.1 points per CCS+LRS category. mRMDQ scores demonstrated smaller changes across all follow-up time points, with decreases ranging from 1.0 to 3.5 points per one-category increase in stenosis severity at 8, 26 and 52 weeks.

Table 4 Bivariate correlations between stenosis scores and clinical parameters

		CCS			LRS			CCS+LRS		
	Follow-up	N	Spearman's ρ	P value	N	Spearman's ρ	P value	N	Spearman's ρ	P value
VAS leg pain	Baseline	152	-0.038	0.646	152	-0.083	0.307	152	-0.021	0.800
	8w	150	-0.144	0.080	150	-0.088	0.285	150	-0.138	0.091
	26w	148	-0.250	0.002*	148	-0.066	0.429	148	-0.184	0.025*
	52w	151	-0.284	< 0.001*	151	-0.129	0.115	151	-0.238	0.003*
VAS back pain	Baseline	152	0.020	0.809	152	-0.031	0.702	152	-0.020	0.805
	8w	149	-0.161	0.051	149	-0.029	0.724	149	-0.065	0.433
	26w	148	-0.280	< 0.001*	148	-0.123	0.135	148	-0.179	0.030*
	52w	151	-0.266	< 0.001*	151	-0.180	0.027*	151	-0.187	0.021*
mRMDQ	Baseline	142	0.036	0.670	142	0.088	0.297	142	0.076	0.369
	8w	133	-0.227	0.009*	133	-0.229	0.008*	133	-0.236	0.006*
	26w	133	-0.194	0.025*	133	-0.190	0.028*	133	-0.144	0.099*
	52w	136	-0.244	0.004*	136	-0.193	0.024*	136	-0.148	0.085

CCS: central canal stenosis; FA: facet arthropathy; FS: foraminal stenosis; LRS: lateral recess stenosis. * P value < 0.05

Table 5 Linear regression analysis with covariates age, gender, BMI, duration of symptoms, HADS at baseline and treatment group.

		CCS			LRS			CCS+LRS		
	Follow-up	B	95% CI	P value	B	95% CI	P value	B	95% CI	P value
VAS leg pain	Baseline	-0.356	-3.111 – 2.399	0.799	--3.766	-12.182 – 4.649	0.378	-0.151	-3.994 – 3.693	0.938
	8w	-3.044	-6.272 – 0.184	0.064	-4.627	-14.564 – 5.309	0.359	-3.919	-1.709 - -8.454	0.090
	26w	-5.587	-11.108 – -0.066	0.047*	-1.237	-18.135 – 15.660	0.885	-4.564	-12.318 – 3.190	0.246
	52w	-7.879	- 11.507 – -4.251	< 0.001*	- 10.773	- 22.469 – 0.924	0.071	- 9.639	- 14.783 – -4.495	< 0.001*
VAS back pain	Baseline	0.809	-2.254 – 3.873	0.602	0.359	-9.060 – 9.779	0.940	0.854	-3.430 – 5.137	0.694
	8w	-2.177	-5.558 – 1.204	0.205	-2.017	-12.537 – -8.504	0.705	-1.807	-6.601 – 2.987	0.457
	26w	-5.818	-9.346 – -2.291	0.001*	-7.603	-18.613 – 3.408	0.174	-6.271	-11.264 – -1.278	0.014*
	52w	-6.488	- 9.924 – -3.051	< 0.001*	-14.332	-25.152 – -3.513	0.010*	-8.066	-12.924 – -3.207	0.001*
mRMDQ	Baseline	0.282	-0.315 – 0.879	0.352	0.854	-0.935 – 2.643	0.347	0.406	-0.425 – 1.237	0.335
	8w	-1.294	-2.106 – -0.483	0.002*	-3.484	-5.926 – 1.043	0.006*	- 1.878	-3.019 – -0.737	0.001*
	26w	-1.081	-1.925 – -0.237	0.013*	-3.244	-5.735 – -0.752	0.011*	-1.468	-2.645 – -0.290	0.015*
	52w	- 0.973	-1.782 – -0.165	0.019*	-2.332	-4.791 – 0.127	0.063	-1.015	-2.161 – 0.131	0.082

CCS: central canal stenosis; FA: facet arthropathy; FS: foraminal stenosis; LRS: lateral recess stenosis. * P value < 0.05

DISCUSSION

This study provides external validation of the LSS grading system proposed by Miskin et al. [13]. Our findings demonstrate substantial inter-reader agreement for CCS and LRS, whereas agreement was slight for FS and fair for FA. These findings are supported by the proportions of exact and within-one-grade agreement. The relatively high agreement for LRS is partly expected due to its three-grade scale, which inherently limits the potential for disagreement. The weighted kappa values for CCS were comparable to those reported by Miskin et al., while agreement for LRS was higher but lower for FS and FA. Overall, the grading system appears reliable for CCS and LRS, but less so for FS and FA.

The lower agreement for FS and FA observed in this study may be attributed to differences in patient populations. Miskin et al. included patients with isolated low back pain, a condition often associated with more pronounced facet degeneration, whereas our cohort consisted of patients with intermittent neurogenic claudication, in whom FA may be less severe or more heterogeneous. Similarly, most patients in our study demonstrated no or mild FS, and fewer than 10% exhibited severe FS, likely due to exclusion of patients requiring fusion surgery in the FELIX trial. Our larger sample size may also have contributed to greater variability, thereby reducing inter-reader agreement. Sensitivity analysis demonstrated improved agreement for FA when grades were collapsed, suggesting that the current FA scale may be overly granular.

Over the years, multiple grading systems have been proposed to assess one or more components of LSS. The most commonly used CCS grading systems are those proposed by Guen et al. and Schizas et al. [3, 4]. Guen et al. reported kappa values of 0.730-0.953, although external validation yielded substantially lower agreement (kappa 0.33-0.40) [18]. Schizas et al. reported a kappa of 0.44, with subsequent studies demonstrating variable inter-reader agreement [19-22]. Two additional CCS grading systems reported kappa values of 0.58-0.76 and 0.681, respectively [5, 6]. For FS, three grading systems have been introduced. Wildermuth et al. reported substantial agreement (kappa 0.62), Lee et al. demonstrated near-perfect agreement (kappa 0.909-0.942), which was externally validated [22], and Sartoretti et al. reported similarly high agreement (kappa 0.886-1.000) [7-9]. For LRS, Pfirrmann et al. proposed a grading scale with substantial agreement (kappa 0.62-0.67), although the cohort consisted of patients with lumbar disc herniation [10]. For FA, two grading scales have been presented. Weishaupt et al. reported a kappa of 0.41 [11], although this was lower in subsequent studies [19, 22], whereas Fujiwara et al. demonstrated a kappa of 0.636 [12], which was substantiated by another study [20]. While some

grading systems demonstrate high inter-reader agreement, most assess only a single anatomical component of LSS, thereby limiting overall clinical utility. Only the system proposed by Lurie et al., similar in concept to that of Miskin et al., incorporates multiple anatomical zones by estimating the proportion of compromised area (normal, ≤1/3, 1/3-2/3, ≥2/3) in the central canal, lateral recess, and neuroforamen. Lurie et al. reported kappa values of 0.73 for CCS, 0.49 for FS, and 0.58 for LRS [2]. In this context, our findings indicate that the Miskin et al. grading system performs comparably to, or better than, most existing grading systems for CCS and LRS, supporting its clinical utility for these components. However, the low agreement observed for FS and FA suggests that these components may require refinement or substitution with alternative, more reliable grading systems.

In the present study, higher baseline stenosis severity, particularly CCS, was independently associated with greater postoperative improvement in pain and physical disability. Although effect sizes per category were modest and below commonly accepted minimally clinical important difference (MCID) thresholds for VAS and RMDQ scores[23], cumulative differences across severity categories approached or exceeded MCID values, indicating potential clinical relevance. These findings were not attributable to a greater capacity for symptomatic improvement in patients with more severe stenosis, as baseline pain and disability did not differ significantly across stenosis severity categories. Consequently, patients with higher baseline CCS severity, as graded using the Miskin et al. scale, appear more likely to benefit from surgery. Importantly, the combined CCS+LRS score did not demonstrate stronger correlations with clinical outcomes than CCS alone.

In contrast, many previously proposed grading systems have not demonstrated meaningful correlations with clinical parameters. Prior studies assessing the Schizas or Weishaupt grading scales reported no significant associations with baseline symptom severity or postoperative outcomes [3, 19, 20], with the exception of one study demonstrating a correlation between Schizas grades and both baseline and postoperative leg pain scores [21]. Yuan et al. reported that their CCS grading system correlated with baseline physical disability but not with postoperative clinical outcomes [6]. Hence, these findings underscore the limited value of many existing grading systems in clinical decision-making.

Strengths and limitations

Strengths of this study include a representative patient population with intermittent neurogenic claudication, a large sample size, variability in MRI scanning parameters, and independent grading by three experienced and blinded review-

ers. Limitations include the exclusion of patients with missing data, potentially introducing selection bias, and the absence of a direct comparison with other grading systems.

Recommendations for future research

While the grading system proposed by Miskin et al. offers a valuable step toward comprehensive LSS assessment, alternative systems, specifically for FS and FA, should not be discounted. A modular approach in which the most reliable grading method is selected for each anatomical region, may be preferable. Accordingly, the grading systems proposed by Miskin et al. for CCS, Lee et al. and Sartoretti et al. for FS, and Fujiwara et al. for FA warrant further external validation. Ultimately, grading systems should be evaluated not only for reproducibility but also for their association with clinical outcomes across all lumbar levels.

Conclusion

The Miskin et al. grading system appears promising for LSS evaluation and clinical decision-making. The substantial interobserver agreement for CCS and LRS, which is comparable to or exceeds that of other grading systems, supports their use in clinical practice and interdisciplinary communication. Moreover, CCS severity correlated with postoperative clinical outcomes and may therefore aid in preoperative counseling. However, limited inter-reader agreement for FS and FA reduces their reliability, and substitution of these components with alternative, more reliable grading systems may improve the overall clinical utility of this approach.

REFERENCES

[1] Andreisek G, Hodler J, Steurer J. Uncertainties in the diagnosis of lumbar spinal stenosis. Radiology. 2011;261(3):681–4.

[2] Lurie JD, Tosteson AN, Tosteson TD, Carragee E, Carrino JA, Kaiser J, et al. Reliability of readings of magnetic resonance imaging features of lumbar spinal stenosis. Spine (Phila Pa 1976). 2008;33(14):1605–10.

[3] Schizas C, Theumann N, Burn A, Tansey R, Wardlaw D, Smith FW, et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphology of the dural sac on magnetic resonance images. Spine (Phila Pa 1976). 2010;35(21):1919–24.

[4] Guen YL, Joon WL, Hee SC, Kyoung-Jin O, Heung SK. A new grading system of lumbar central canal stenosis on MRI: an easy and reliable method. Skeletal Radiology. 2011;40(8):1033–9.

[5] Weber C, Rao V, Gulati S, Kvistad KA, Nygaard Ø P, Lønne G. Inter- and Intraobserver Agreement of Morphological Grading for Central Lumbar Spinal Stenosis on Magnetic Resonance Imaging. Global Spine J. 2015;5(5):406–10.

[6] Yuan S, Zou Y, Li Y, Chen M, Yue Y. A clinically relevant MRI grading system for lumbar central canal stenosis. Clin Imaging. 2016;40(6):1140–5.

[7] Wildermuth S, Zanetti M, Duewell S, Schmid MR, Romanowski B, Benini A, et al. Lumbar spine: quantitative and qualitative assessment of positional (upright flexion and extension) MR imaging and myelography. Radiology. 1998;207(2):391–8.

[8] Lee S, Lee JW, Yeom JS, Kim KJ, Kim HJ, Chung SK, et al. A practical MRI grading system for lumbar foraminal stenosis. AJR Am J Roentgenol. 2010;194(4):1095–8.

[9] Sartoretti E, Wyss M, Alfieri A, Binkert CA, Erne C, Sartoretti-Schefer S, et al. Introduction and reproducibility of an updated practical grading system for lumbar foraminal stenosis based on high-resolution MR imaging. Sci Rep. 2021;11(1):12000.

[10] Pfirrmann CW, Dora C, Schmid MR, Zanetti M, Hodler J, Boos N. MR image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. Radiology. 2004;230(2):583–8.

[11] Weishaupt D, Zanetti M, Boos N, Hodler J. MR imaging and CT in osteoarthritis of the lumbar facet joints. Skeletal Radiol. 1999;28(4):215–9.

[12] Fujiwara A, Tamai K, Yamato M, An HS, Yoshida H, Saotome K, et al. The relationship between facet joint osteoarthritis and disc degeneration of the lumbar spine: an MRI study. Eur Spine J. 1999;8(5):396–401.

[13] Miskin N, Isaac Z, Lu Y, Makhni MC, Sarno DL, Smith TR, et al. Simplified Universal Grading of Lumbar Spine MRI Degenerative Findings: Inter-Reader Agreement of Non-Radiologist Spine Experts. Pain Med. 2021;22(7):1485–95.

[14] Moojen WA, Arts MP, Brand R, Koes BW, Peul WC. The Felix-trial. Double-blind randomization of interspinous implant or bony decompression for treatment of spinal stenosis related intermittent neurogenic claudication. BMC Musculoskelet Disord. 2010;11:100.

[15] Zanoli G, Strömqvist B, Jönsson B. Visual analog scales for interpretation of back and leg pain intensity in patients operated for degenerative lumbar spine disorders. Spine (Phila Pa 1976). 2001;26(21):2375–80.

[16] Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine (Phila Pa 1976). 1995;20(17):1899–908; discussion 909.

[17] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.
[18] Al-Tameemi HN, Al-Essawi S, Shukri M, Naji FK. Using Magnetic Resonance Myelography to Improve Interobserver Agreement in the Evaluation of Lumbar Spinal Canal Stenosis and Root Compression. Asian Spine J. 2017;11(2):198–203.
[19] Moojen WA, Schenck CD, Lycklama À Nijeholt GJ, Jacobs WCH, Van der Kallen BF, Arts MP, et al. Preoperative MRI in Patients With Intermittent Neurogenic Claudication: Relevance for Diagnosis and Prognosis. Spine (Phila Pa 1976). 2018;43(5):348–55.
[20] Minetama M, Kawakami M, Teraguchi M, Matsuo S, Enyo Y, Nakagawa M, et al. MRI grading of spinal stenosis is not associated with the severity of low back pain in patients with lumbar spinal stenosis. BMC Musculoskelet Disord. 2022;23(1):857.
[21] Yang J, Xiong Y, Hu Y, Huang M, Zhang L, Pu X, et al. The reliability, correlation with clinical symptoms and surgical outcomes of dural sac cross-sectional area, nerve root sedimentation sign and morphological grade for lumbar spinal stenosis. BMC Musculoskelet Disord. 2023;24(1):225.
[22] Banitalebi H, Espeland A, Anvar M, Hermansen E, Hellum C, Brox JI, et al. Reliability of preoperative MRI findings in patients with lumbar spinal stenosis. BMC Musculoskelet Disord. 2022;23(1):51.
[23] Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976). 2008;33(1):90–4.

10

Artificial intelligence for segmentation and classification in lumbar spinal stenosis: an overview of current methods

E.J.A. Verheijen, T. Kapogiannis, D. Munteh, J. Chabros,
M. Staring, T.R. Smith, C.L.A. Vleggeert-Lankamp

Eur Spine J 2025

ABSTRACT

Purpose

Lumbar spinal stenosis (LSS) is a frequently occurring condition defined by narrowing of the spinal or nerve root canal due to degenerative changes. Physicians use MRI scans to determine the severity of stenosis, occasionally complementing it with X-ray or CT scans during the diagnostic work-up. However, manual grading of stenosis is time-consuming and induces inter-reader variability as a standardized grading system is lacking. Machine Learning (ML) has the potential to aid physicians in this process by automating segmentation and classification of LSS. However, it is unclear what models currently exist to perform these tasks.

Methods

A systematic review of literature was performed by searching the Cochrane Library, Embase, Emcare, PubMed, and Web of Science databases for studies describing an ML-based algorithm to perform segmentation or classification of the lumbar spine for LSS. Risk of bias was assessed through an adjusted version of the Newcastle-Ottawa Quality Assessment Scale that was more applicable to ML studies. Qualitative analyses were performed based on type of algorithm (conventional ML or Deep Learning (DL)) and task (segmentation or classification).

Results

A total of 27 articles were included of which nine on segmentation, 16 on classification and 2 on both tasks. The majority of studies focused on algorithms for MRI analysis. There was wide variety among the outcome measures used to express model performance. Overall, ML algorithms are able to perform segmentation and classification tasks excellently. DL methods tend to demonstrate better performance than conventional ML models. For segmentation the best performing DL models were U-Net based. For classification U-Net and unspecified CNNs powered the models that performed the best for the majority of outcome metrics. The number of models with external validation was limited.

Conclusion

DL models achieve excellent performance for segmentation and classification tasks for LSS, outperforming conventional ML algorithms. However, comparisons between studies are challenging due to the variety in outcome measures and test datasets. Future studies should focus on the classification task using DL models and utilize a standardized set of outcome measures and publicly available test dataset to express model performance. In addition, these models need to be externally validated to assess generalizability.

INTRODUCTION

Lumbar spinal stenosis (LSS) is a disease defined by narrowing of the spinal or nerve root canal that becomes symptomatic through the compression of neural structures [1]. Classically, LSS causes intermittent neurogenic claudication affecting approximately 11% of older adults in the US and is the most common cause of spinal surgery among this population [2,3,4].

Patients are usually offered surgery if conservative treatment has failed to sufficiently ameliorate symptoms. Surgical candidacy is assessed using radiological imaging in conjunction with clinical history and physical examination [4, 5]. MRI has become the gold standard to determine the severity of stenosis, as it produces detailed images of relevant soft tissues that may contribute to the stenosis [6, 7]. Alternatively, Computed Tomography (CT) may be used in assessing the bony component of stenosis, although delivering less valuable information on true compression of the cauda equina or spinal nerve root. Evaluation of the severity of stenosis can be subjective, with inter-reader variability among radiologists and surgeons [8, 9]. Grading systems to standardize MRI interpretation for the severity of stenosis, such as those proposed by Lee et al. [10], Schizas et al. [11], and Miskin et al. [12], have demonstrated inter-observer metrics ranging from "fair" to "excellent reliability" (Cohen's kappa 0.323–0.702, intraclass correlation coefficient 0.730–0.953), and, hence, have not eliminated variability. In addition, manually assessing MRIs using these grading systems is time-consuming and, thus, not feasible in clinical practice.

Artificial Intelligence (AI) may be a valuable tool to assist clinicians in grading LSS, as it has demonstrated the ability to assess medical images accurately and consistently in other disease areas [13]. Conventional machine learning (ML) or deep learning (DL) architectures can be trained for image analysis either through supervised or through unsupervised learning. In supervised learning, training images are labeled, and this technique is often used for segmentation and outcome prediction [14]. In practice, semi-supervised and weakly-supervised approaches are more common, especially when high-quality labeled data is scarce. Semi-supervised learning combines a limited amount of labeled data with a large amount of unlabeled data, whereas weakly-supervised learning relies on imperfect or imprecise labels when accurate labeling is challenging or costly. In contrast, unsupervised learning models can detect patterns in unlabeled data, which is valuable as image datasets with high-quality labeling are difficult to procure [13]. The conventional ML approach to image analysis usually necessitates the selection of relevant input features to train the algorithm to complete two tasks: segmentation and classification. In segmentation, each

pixel is assigned to a class based on its extracted attributes (Fig. 1) and are then used as inputs for classification, where predictions are generated on the severity of LSS [15]. DL models are a subset of ML that can learn important features from the raw data, obviating the need for extensive feature engineering, do not require the segmentation step before classification and have demonstrated stronger performance than conventional ML before [16, 17] (Fig. 2).

In this systematic review, we describe current conventional ML and DL models for segmentation and classification of LSS, including scoring of their performance. We aim to examine whether AI can be used to improve LSS diagnostics.

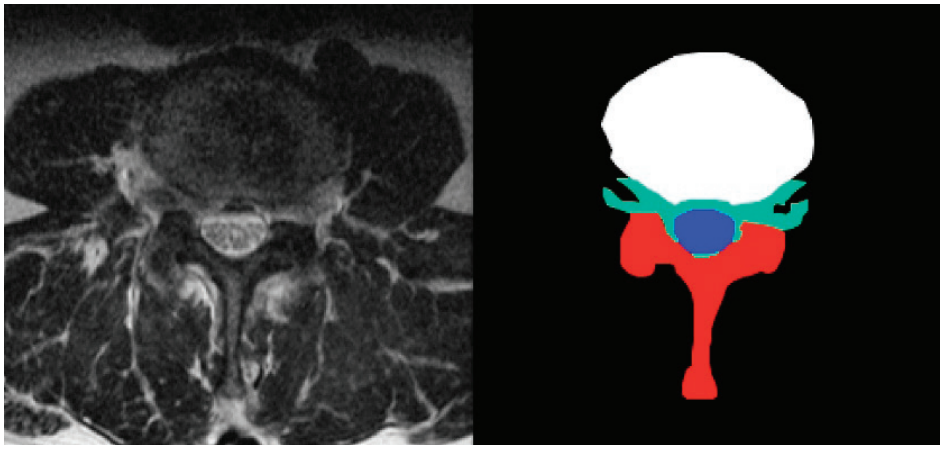


Fig. 1 Example of segmentation of spinal structures in the lumbar spine from an axial T2-weighted MRI image

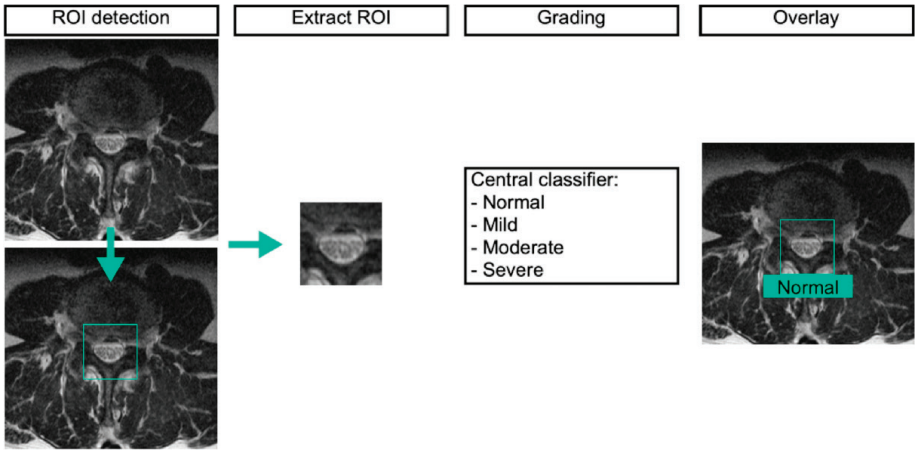


Fig. 2 Example of classification of central spinal stenosis. The algorithm determines the region of interest (ROI) on an axial T2-weighted MRI image, extracts it and decides on the grading of stenosis

METHODS

This systematic review was conducted in accordance with the PRISMA guidelines.

Search and selection

Relevant articles were searched in five databases (Cochrane Library, Embase, Emcare, PubMed, and Web of Science) from inception to 9 February 2023. An expert librarian created a comprehensive search strategy that included strings for studies investigating new or validating existing algorithms for segmentation or classification of LSS (Online Resource 1). All search results were screened by two reviewers (EV and JC) separately based on title and abstract. The remaining full texts were evaluated, and, consequently, screening of references and citation tracking were performed. Any discrepancies were resolved by discussion or consulting a third reviewer (CVL).

Inclusion and exclusion criteria

Studies describing segmentation or classification algorithms for LSS in adults based on conventional ML or DL approaches were considered for eligibility. Segmentation was defined as the ability to label spinal structures on a pixel-level. The placement of a bounding box to extract a region of interest (ROI) was not considered segmentation. For classification, studies were accepted that categorized patients according to severity of stenosis (either binary or multiple classes) or that automated measuring spine indices relevant for classification of LSS. Meeting abstracts, case reports, systematic reviews and meta-analyses were excluded. Only algorithms developed for routine X-ray, CT or MRI were accepted (e.g., excluding non-routine MRI myelography scans). In addition, segmentation studies were required to assess at least two anatomical structures relevant for LSS (e.g., IVD, spinal canal, lateral recess), since LSS is considered a multifactorial disorder. This was also a criterion for classification studies that did not directly classify the degree of stenosis but rather, e.g., degree of disc degeneration and hypertrophy of the ligaments. Only articles written in English and available in full text were included. If an article described a continuation or improvement of previous work by the same author(s), only the most complete work was included.

The aim is to develop a model with the ability to correctly identify and predict classes (discriminate between two or more conditions) with high (spatial/geometric) accuracy and consistent with expert knowledge. Therefore, we defined acceptable outcome measures as those quantifying accuracy (e.g., overall accuracy, F1 score, area under the curve (AUC)), spatial/geometric reliability/error

(e.g., area, Hausdorff distance), and similarity coefficients quantifying agreement with ground truth (e.g., Cohen's kappa, Jaccard index, Dice coefficient).

Risk-of-bias assessment

Risk of bias was assessed by three reviewers separately (EV, TK and DM) using an adjusted version of the Newcastle–Ottawa Quality Assessment Scale [18]. At the time of assessment, an AI-specific risk-of-bias tool had not been developed, although the authors were aware of the work in progress on this matter[19]. The risk-of-bias criteria were adopted to better fit our research aim and be more applicable to studies evaluating computer algorithms (Online Resource 2). Each study could be awarded 0 to 10 points. Studies with a score above seven points were classified as low risk of bias, studies with 5–7 points as intermediate risk of bias and studies with fewer than 5 points as high risk of bias. Differences between the reviewers were resolved during a consensus meeting or with a fourth reviewer (CVL).

Data extraction and analysis

From all included studies, data was collected by two reviewers (TK and DM) on: year of publication, radiological scans (modality, slice orientation, scanning parameters and field strength if applicable), algorithm (type of model, architecture, use of transfer learning, type of loss function and optimization, degree of automation), study comparisons, ground truth, data handling (sample size, preprocessing, augmentation, imbalance, split), validation and testing, outcome measures and results. A third reviewer (EV) verified the final data extraction sheet. In cases where authors presented results for different versions of the same model architecture (e.g., different loss function, varying hyperparameters) or with different thresholds (e.g., varying distance error tolerance), only data were collected for the average of the models or, if the average was not provided, the best performing algorithm, or where the strictest outcome criteria were applied. The heterogeneity among outcome measures precluded pooling of the results, and, therefore, a qualitative analysis was performed. Articles were compared in four categories: conventional ML segmentation, DL segmentation, conventional ML classification and DL classification. Within a category, results were compared that belonged to the same type of outcome measure (i.e., measure of accuracy, spatial/geometric reliability/error or similarity metrics).

RESULTS

Article selection

The initial literature search yielded a total of 661 unique articles. Of those, 616 were excluded and the remaining 29 articles were selected for full-text screening of which 22 were included. After citation tracking an additional 16 studies were screened in full text of which five were accepted. Ultimately, 27 studies were included for this review (Fig. 3) [20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46]. A comprehensive overview of the included studies is provided in Online Resource 3.

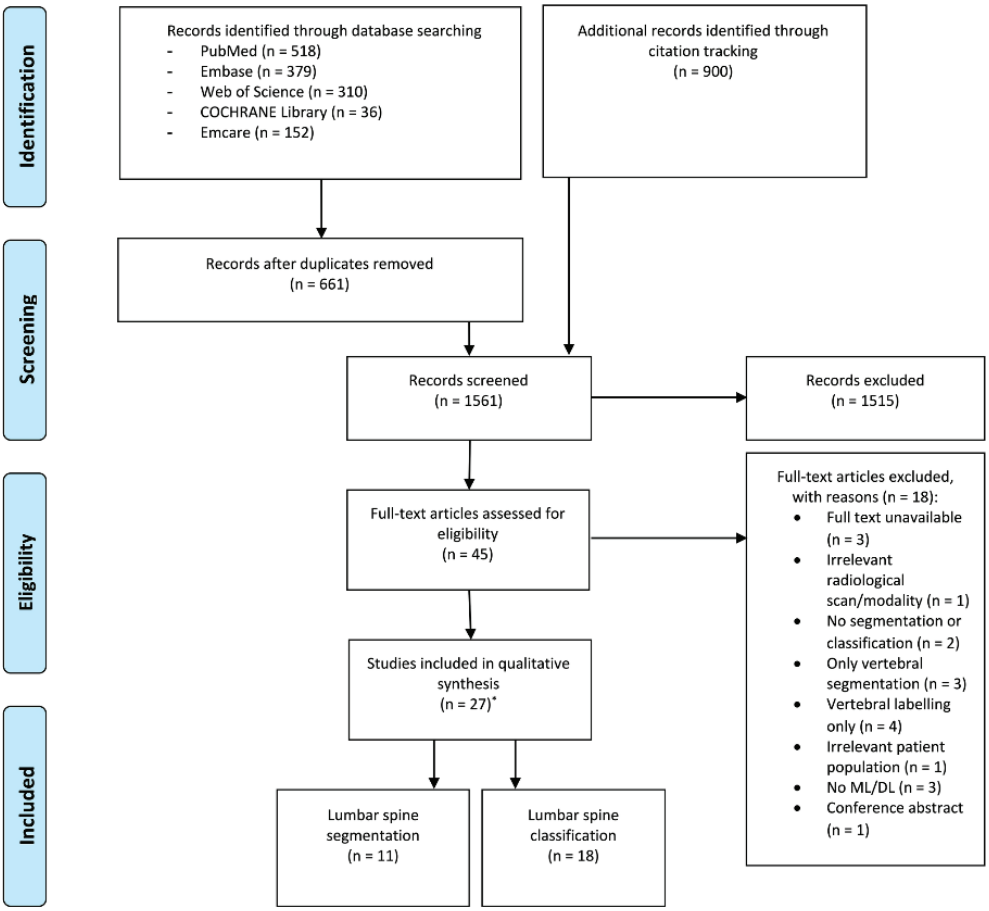


Fig. 3 Flowchart of the article search and selection process.

* Two studies reported on both segmentation and classification and, therefore, appear in both segmentation and classification boxes

Nine articles assessed segmentation, 16 articles assessed classification and two reported on both. Year of publication ranged between 2010 and 2023 with most of the studies being published in 2020 or thereafter. Of the 20 studies published between 2019 and 2023, seven reported on segmentation and 13 reported on classification, whereas, of the six studies published between 2014 and 2018, four reported on segmentation and three reported on classification, demonstrating a shift of focus towards classification challenges (Fig. 4).

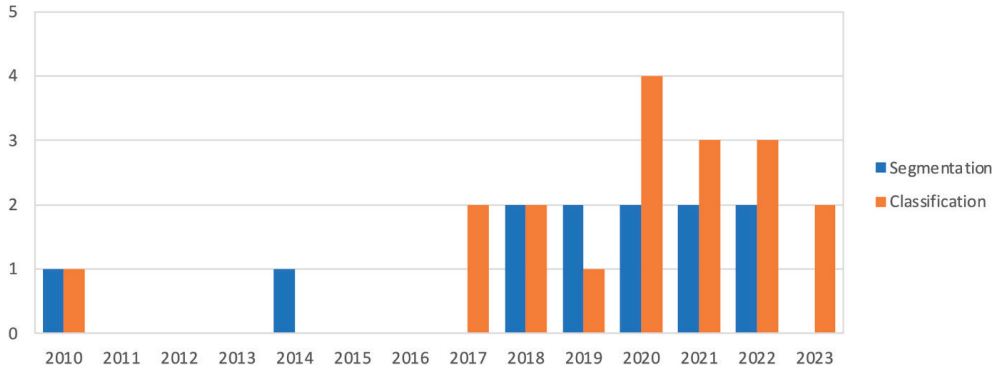


Fig. 4 Number of publications per year focusing on ML/DL algorithms for segmentation or classification of the lumbar spine for LSS. Publications that assessed both appear twice in this graph

Risk-of-bias assessment

Of the 27 studies, three were considered low-risk of bias [26, 27, 31], nineteen were categorized as intermediate-risk of bias [20,21,22, 24, 28, 30, 32,33, 34, 36,37,38,39,40,41,42,43,44, 46] and five were judged as high-risk of bias [23, 25, 29, 35, 45] (Online Resource 4). Risk of bias was higher for segmentation than classification studies on average, but independent of year of publication.

Segmentation algorithms

Segmentation of spinal structures included the IVD, spinal canal, thecal sac (TS), posterior element (PE)/lamina, ligamentum flavum, facet joints, neural foramina, and the area between anterior and posterior vertebrae elements (AAP). Two out of eleven studies employed conventional ML methods, eight used DL, and one applied both techniques. The extracted data are provided in Online Resource 5 and 6.

Conventional machine learning for segmentation

A variety of outcome measures were used to express model performance. None used accuracy, but one study described precision (0.79–0.83) and sensitivity

(0.90–0.92) for disc and dural sac segmentation [25] and another study devised an own metric to express segmentation quality (91.25–98.21%) [35]. Only one study presented spatial metrics reporting Hausdorff distances (7.89–9.41 mm) and surface distances (0.83–0.84) [24]. Two studies presented similarity metrics: one study demonstrated Dice scores ranging 0.83–0.84 [24], while another reported Dice scores of 0.84–0.87 and a Jaccard index between 0.73 and 0.78 [25]. Since Koompairojn et al. received a relatively high risk-of-bias score, the model by Ghosh et al. was considered most reliable [25]. This fully automatic model comprised Histogram of Oriented Gradients (HOG) feature descriptors and random forests (RF) as the classifier with a fixed number of trees (n=100) based on T2-weighted sagittal MRI scans from 50 patients. They used fivefold cross validation and tested on 212 patients.

Deep learning for segmentation

Eight studies performed segmentation using MRI [20, 21, 23, 24, 28, 29, 39, 42] and one used CT scans [45]. Five studies were U-Net based models [23, 24, 29, 39, 45], two studies were SegNet based [20, 42], one study was Generative Adversarial Network (GAN) based [28] and one study tested a standard convolutional neural network (CNN) [21].

An overall accuracy of 85% was reported in one study [47], but a range between 50 and 99% for different spinal structures in another [20]. Four studies provided pixel accuracy with Hou et al. demonstrating the highest accuracy (0.9935) [29]. Other accuracy metrics were only measured in a single study. Only one article reported a geometrical metric achieving a surface distance of 2.71 mm [39]. The most commonly reported performance metrics were similarity coefficients. The Jaccard index was used in five reports with a best overall score of 0.8493 [45], although higher scores were demonstrated for specific spinal structures [20]. When correcting for class frequency, an even higher score was achieved (0.9835) [45]. The Dice coefficient was reported in four studies with a maximum value of 0.9252 [39]. Due to the variety in reported performance metrics it was not possible to determine a single most reliable DL model for segmentation. However, considering (pixel) accuracy, Jaccard index and Dice score, the best performing models were U-Net based [29, 39, 45] or SegNet based [20]. All four compared their model's performance to human expertise using data of at least 120 patients and evaluated their own model (SegNet-TL80 [20], Spine-Seg-2 phase model [29], MANet [39] and DDU-Net [45]. Two were semi-automated [29, 39] and two were fully automated [20, 45].

Classification algorithms

A total of eighteen articles assessing classification algorithms for LSS were published between 2010 and 2023 [22, 26,27,28, 30,31,32,33,34,35,36,37,38, 40, 41, 43, 44, 46]. Two out of eighteen studies described conventional ML algorithms, fifteen used DL and one compared both techniques. The extracted data are provided in Online Resource 7 and 8.

Conventional machine learning for classification

Of the three studies, Koompairojn et al. demonstrated an overall accuracy of 92.66–96.82% for different spinal features [35]. Altun et al. achieved the highest accuracy (0.762) with a Gabor-RF model using one axial image for each of three IVD levels (L3-4 to L5-S1) [22]. Huber et al. evaluated performance through sensitivity, specificity and AUC, and obtained superior results with a decision tree algorithm that used 240 texture analysis features from the dural sac or spinal canal as input (sensitivity: 94.32–94.33%, specificity: 96.53–98.04%, AUC: 0.940–0.962) [30]. The ground truth was established using a reference standard for the cross-sectional area, defined as above 130 mm^2 (non-severe stenosis) or below (severe stenosis). A total of 343 images from 82 patients (max. 5 axial slices per patient, one per IVD level) were used with tenfold cross validation. This model was considered semi-automated since it required manually segmented MRI images.

Deep learning for classification

Sixteen studies described DL models for classification [22, 26,27,28, 31,32,33,34, 36,37,38, 40, 41, 43, 44, 46]. Six studies employed an unspecified CNN [26, 27, 31,32,33, 40], five studies were VGG-based [22, 34, 37, 38, 46], two studies were U-Net based [36, 43], two studies were ResNet-based [41, 44], and one study used a GAN [28]. Models either classified specific spinal features with binary (stenosis or not) or multiclass labels (e.g., normal, mild, moderate, severe stenosis), performed measurements between spinal structures, or classified the image as a whole (binary label).

Out of six studies reporting on accuracy, Lehnen et al. achieved the highest score (98.09%) [36]. Class average accuracy was reported in four studies, with the highest range of values achieved by Jamaludin et al. (0.701–0.947) [32]. AUC was reported the highest by Lu et al. (0.961–0.983) among five studies [41]. Sensitivity and specificity were presented in 9 studies: the highest overall sensitivity was achieved by Altun et al. (0.921) [22], although two studies reported higher values for central canal stenosis specifically (0.922 and 0.946) [27, 44].

Lehnen et al. reported the highest specificity (0.9865) [36]. Negative (NPV) and positive predictive value (PPV) were only reported in three and five studies, respectively. Lehnen et al. achieved an NPV of 0.9906, whereas PPV was highest in the study by Lewandrowski et al. [38], although Kim et al. achieved values between 0.790 and 0.845 depending on the use of neutral, flexion or extension radiographs [34]. F1 score was highest in the study by Su et al., although this was only reported by one other study [44]. Among spatial metrics only the mean absolute error (MAE) for the diameter of the dural sac was reported in two studies with the smallest error obtained by Pang et al. (0.72 mm) [40, 43]. For similarity metrics Cohen's kappa was reported in four studies but highest in the study by Ishimoto et al. (0.75) [31]. Gwet's AC1 statistic, reported in two studies, was highest in the study by Hallinan et al. (0.68–0.96) [27]. In addition, Lin's correlation coefficient was presented in three studies with the best performance reported by Jamaludin et al. (0.88–0.89) [32].

Similar to the results for segmentation models, there was a wide variation in the performance metrics for classification. Moreover, some models achieved high scores for one metric, but lower for other outcome measures. As a result, nine different studies achieved the best score depending on the performance metric. Of the twelve metrics, the best result for four parameters was achieved using U-Net based models [36, 43], for four others unspecified CNN algorithms worked best [27, 31, 32], for two ResNet derived models obtained the best results [41, 44] and for another two metrics VGG-like models performed the best [22, 37]. Two studies used multiple axial and sagittal images (axial: 15–28 slices, sagittal: 15 slices [27, 44], three studies used only one image for each IVD level [22, 31, 43], and 4 studies did not specify the number of slices per patient or per IVD level [32, 36, 37, 41]. Hence, it was not possible to identify one DL algorithm as the most reliable model for the classification task.

DISCUSSION

Through this systematic review, an overview is presented of the currently developed conventional ML and DL techniques for segmentation and classification tasks of LSS. It has been demonstrated that there is a wide spectrum of models for these aims, mostly based on MRI scans. In addition, there is a preponderance of DL models among the studies that were published more recently. Although the paucity of comparable performance metrics posed a challenge for comparisons, DL models, especially U-Net based, yielded better results than conventional ML algorithms for the segmentation task. For classification purposes, DL networks were superior to conventional ML solutions, specifically U-Net and

'standard' CNN algorithms performed the best for most of the performance metrics, obviating the need for a preceding segmentation step. Overall, the models demonstrated results nearing human performance which holds promise to support physicians in the diagnosis of LSS.

LSS can occur centrally around the thecal sac, in the lateral recess or at the neural foramen. Therefore, the ideal classification model would include all these structures. Three out of the nine best performing DL classification algorithms labelled the central canal, lateral recess and neural foramen (or lateral recess and neural foramen combined as 'nerve root') for stenosis [27, 36, 44] and one model labelled the whole image [22]. In general, central and lateral recess stenosis are best assessed using axial MRI images, whereas the neural foramina can be most optimally viewed on sagittal scans. Since the studies by Hallinan et al. and Lehnen et al. included both planes, they can be considered the most comprehensive models for classification of LSS [27, 36], and, as a result, hold the most clinical value. However, both models have only been externally validated to a limited extent. Hallinan et al. tested their model on an external dataset of 100 patients from Saudi Arabia following training and testing with patient data from a hospital in Singapore. Lehnen et al. validated the commercially available CoLumbo model originating from Bulgaria with imaging from their institution in Germany. Yet, the generalizability of these models could be further substantiated by additional validation studies.

The heterogeneity among the performance metrics reported precluded a quantitative analysis of the results for both segmentation and classification algorithms. There was no subset of metrics that was consistently reported in every study, creating a challenge for appropriate comparisons between the studies. To overcome this problem in future studies, reporting on ML algorithms needs to be standardized, and recommendations have been made to guide authors in this process [48, 49]. Additionally, more specific guidelines are being developed which may include a combination of performance metrics that should be reported for every ML model [19]. For segmentation algorithms, we propose that accuracy and the spatial error in mm should be reported at a minimum. For classification models, authors should report accuracy, the confusion matrix and area under the receiver operating characteristic curve [50]. Furthermore, a publicly available test dataset should be used to improve comparability between studies and assess generalizability. Ideally, such a dataset would contain multi-vendor MRI images from different medical centers across the world with high-quality labelling. To our knowledge, only one publicly available dataset with corresponding ground truth data is currently available containing axial and sagittal MRI images from 515 patients obtained across several international institu-

tion [51]. It was used by four studies from this review [20, 29, 42, 43], however, this dataset only contains segmentation ground truth data for one axial scan of the last three IVDs [52]. Radiologists' readings reporting on the presence of central or foraminal stenosis are available [53], but, ideally, classification labels should be derived from a consensus amongst spine-dedicated radiologists and surgeons. Furthermore, the scanning parameters are homogenous which does not reflect the heterogeneity of scans in clinical practice, and no demographic information is provided. Other datasets are available but with limited size and without appropriate ground truth data [54]. Furthermore, authors should clearly state the sample sizes and inter- and intra-observer variability when comparing with human performance. Notably, most studies in this review did not provide details on data or code sharing which defies transparency and developments in AI research [49].

This systematic review is limited by the databases that were searched since relevant articles may have been published in non-medical journals. Furthermore, the lack of consistency in reported outcome metrics precluded quantitative analyses.

Future studies should focus on DL models for classification tasks since they outperform conventional ML methods and obviate the need for segmentation. Furthermore, future studies should validate previously published models in addition to developing new ones. Of the 27 articles in this review, only seven discussed external validation [26, 27, 31, 34, 36, 38, 44]. It is essential that published algorithms are tested on external datasets for assessment of generalizability using a standardized set of performance metrics and a publicly available test dataset with high-quality ground truth data. Additionally, authors should consider the implementability of their classification algorithms, as the ultimate goal is to create a useful model for daily practice. Several models required manual pre-processing of the input data, and, therefore, will not be viable for the high volume of imaging in clinical practice.

This systematic review provides an overview of current conventional ML and DL methods for LSS segmentation and classification. We have elucidated the wide range of developed models and the tendency of DL methods to perform better than conventional ML models obviating the need for a segmentation step before classification. It is essential that guidelines are developed for reporting of performance metrics and that researchers focus on validation of (current) models on external datasets. Primarily for classification, DL models have great potential to improve diagnostics and aid clinicians in LSS identification.

Online Resources can be reviewed online:
https://doi.org/10.1007/s00586-025-08672-9

REFERENCES

1. Katz JN, Harris MB. Clinical practice. Lumbar spinal stenosis. N Engl J Med. 2008;358(8):818-25.
2. Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and Management of Lumbar Spinal Stenosis: A Review. JAMA. 2022;327(17):1688-99.
3. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. Jama. 2010;303(13):1259-65.
4. Melancia JL, Francisco AF, Antunes JL. Chapter 35 - Spinal stenosis. In: Biller J, Ferro JM, editors. Handbook of Clinical Neurology. 119: Elsevier; 2014. p. 541-9.
5. Zileli M, Crostelli M, Grimaldi M, Mazza O, Anania C, Fornari M, et al. Natural Course and Diagnosis of Lumbar Spinal Stenosis: WFNS Spine Committee Recommendations. World Neurosurg X. 2020;7:100073.
6. Seo J, Lee JW. Magnetic Resonance Imaging Grading Systems for Central Canal and Neural Foraminal Stenoses of the Lumbar and Cervical Spines With a Focus on the Lee Grading System. Korean J Radiol. 2023;24(3):224-34.
7. Bozzo A, Marcoux J, Radhakrishna M, Pelletier J, Goulet B. The role of magnetic resonance imaging in the management of acute spinal cord injury. J Neurotrauma. 2011;28(8):1401-11.
8. Winklhofer S, Held U, Burgstaller JM, Finkenstaedt T, Bolog N, Ulrich N, et al. Degenerative lumbar spinal canal stenosis: intra- and inter-reader agreement for magnetic resonance imaging parameters. Eur Spine J. 2017;26(2):353-61.
9. Speciale AC, Pietrobon R, Urban CW, Richardson WJ, Helms CA, Major N, et al. Observer variability in assessing lumbar spinal stenosis severity on magnetic resonance imaging and its relation to cross-sectional spinal canal area. Spine (Phila Pa 1976). 2002;27(10):1082-6.
10. Guen YL, Joon WL, Hee SC, Kyoung-Jin O, Heung SK. A new grading system of lumbar central canal stenosis on MRI: an easy and reliable method. Skeletal Radiology. 2011;40(8):1033-9.
11. Schizas C, Theumann N, Burn A, Tansey R, Wardlaw D, Smith FW, et al. Qualitative Grading of Severity of Lumbar Spinal Stenosis Based on the Morphology of the Dural Sac on Magnetic Resonance Images. Spine. 2010;35(21):1919-24.
12. Miskin N, Isaac Z, Lu Y, Makhni MC, Sarno DL, Smith TR, et al. Simplified Universal Grading of Lumbar Spine MRI Degenerative Findings: Inter-Reader Agreement of Non-Radiologist Spine Experts. Pain Medicine. 2021;22(7):1485-95.
13. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer. 2018;18(8):500-10.
14. Aljuaid A, Anwar M. Survey of Supervised Learning for Medical Image Processing. SN Computer Science. 2022;3(4):292.
15. Seo H, Badiei Khuzani M, Vasudevan V, Huang C, Ren H, Xiao R, et al. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. Med Phys. 2020;47(5):e148-e67.
16. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017;42:60-88.

17. Wang H, Zhou Z, Li Y, Chen Z, Lu P, Wang W, et al. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from (18)F-FDG PET/CT images. EJNMMI Res. 2017;7(1):11.
18. Wells G, Shea B, O'Connell J. The Newcastle-Ottawa Scale (NOS) for Assessing The Quality of Nonrandomised Studies in Meta-analyses. Ottawa Health Research Institute Web site. 2014;7.
19. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open. 2021;11(7):e048008.
20. Al-Kafri AS, Sudirman S, Hussain A, Al-Jumeily D, Natalia F, Meidia H, et al. Boundary Delineation of MRI Images for Lumbar Spinal Stenosis Detection Through Semantic Segmentation Using Deep Neural Networks. IEEE Access. 2019;7:43487-501.
21. Al-Kafri AS, Sudirman S, Hussain AJ, Al-Jumeily D, Fergus P, Natalia F, et al., editors. Segmentation of Lumbar Spine MRI Images for Stenosis Detection Using Patch-Based Pixel Classification Neural Network. 2018 IEEE Congress on Evolutionary Computation (CEC); 2018 8-13 July 2018.
22. Altun S, Alkan A, Altun İ. LSS-VGG16: Diagnosis of Lumbar Spinal Stenosis With Deep Learning. Clin Spine Surg. 2023;36(5):E180-e90.
23. Altun S, Alkan A. LSS-net: 3-dimensional segmentation of the spinal canal for the diagnosis of lumbar spinal stenosis. International Journal of Imaging Systems and Technology. 2022;33(1):378-88.
24. Gaonkar B, Villaroman D, Beckett J, Ahn C, Attiah M, Babayan D, et al. Quantitative Analysis of Spinal Canal Areas in the Lumbar Spine: An Imaging Informatics and Machine Learning Study. AJNR Am J Neuroradiol. 2019;40(9):1586-91.
25. Ghosh S, Chaudhary V. Supervised methods for detection and segmentation of tissues in clinical lumbar MRI. Computerized Medical Imaging and Graphics. 2014;38(7):639-49.
26. Grob A, Loibl M, Jamaludin A, Winklhofer S, Fairbank JCT, Fekete T, et al. External validation of the deep learning system "SpineNet" for grading radiological features of degeneration on MRIs of the lumbar spine. Eur Spine J. 2022;31(8):2137-48.
27. Hallinan J, Zhu L, Yang K, Makmur A, Algazwi DAR, Thian YL, et al. Deep Learning Model for Automated Detection and Classification of Central Canal, Lateral Recess, and Neural Foraminal Stenosis at Lumbar Spine MRI. Radiology. 2021;300(1):130-8.
28. Han Z, Wei B, Mercado A, Leung S, Li S. Spine-GAN: Semantic segmentation of multiple spinal structures. Med Image Anal. 2018;50:23-35.
29. Hou C, Zhang W, Wang H, Liu F, Liu D, Chang J. A semantic segmentation model for lumbar MRI images using divergence loss. Applied Intelligence. 2023;53(10):12063-76.
30. Huber FA, Stutz S, Vittoria de Martini I, Mannil M, Becker AS, Winklhofer S, et al. Qualitative versus quantitative lumbar spinal stenosis grading by machine learning supported texture analysis-Experience from the LSOS study cohort. Eur J Radiol. 2019;114:45-50.
31. Ishimoto Y, Jamaludin A, Cooper C, Walker-Bone K, Yamada H, Hashizume H, et al. Could automated machine-learned MRI grading aid epidemiological studies of

lumbar spinal stenosis? Validation within the Wakayama spine study. BMC Musculoskelet Disord. 2020;21(1):158.
32. Jamaludin A, Lootus M, Kadir T, Zisserman A, Urban J, Battié MC, et al. ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. Eur Spine J. 2017;26(5):1374-83.
33. Jamaludin A, Kadir T, Zisserman A. SpineNet: Automated classification and evidence visualization in spinal MRIs. Med Image Anal. 2017;41:63-73.
34. Kim T, Kim YG, Park S, Lee JK, Lee CH, Hyun SJ, et al. Diagnostic triage in patients with central lumbar spinal stenosis using a deep learning system of radiographs. J Neurosurg Spine. 2022:1-8.
35. Koompairojn S, Hua K, Hua K, Srisomboon J. Computer-Aided Diagnosis of Lumbar Stenosis Conditions. Proc SPIE. 2010.
36. Lehnen NC, Haase R, Faber J, Rüber T, Vatter H, Radbruch A, et al. Detection of Degenerative Changes on MR Images of the Lumbar Spine with a Convolutional Neural Network: A Feasibility Study. Diagnostics (Basel). 2021;11(5).
37. Lewandrowsk IK, Muraleedharan N, Eddy SA, Sobti V, Reece BD, Ramírez León JF, et al. Feasibility of Deep Learning Algorithms for Reporting in Routine Spine Magnetic Resonance Imaging. Int J Spine Surg. 2020;14(s3):S86-s97.
38. Lewandrowski KU, Muraleedharan N, Eddy SA, Sobti V, Reece BD, Ramírez León JF, et al. Reliability Analysis of Deep Learning Algorithms for Reporting of Routine Lumbar MRI Scans. Int J Spine Surg. 2020;14(s3):S98-s107.
39. Li H, Luo H, Huan W, Shi Z, Yan C, Wang L, et al. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network. Neural Comput Appl. 2021;33(18):11589-602.
40. Lin L, Tao X, Yang W, Pang S, Su Z, Lu H, et al. Quantifying Axial Spine Images Using Object-Specific Bi-Path Network. IEEE Journal of Biomedical and Health Informatics. 2021;25(8):2978-87.
41. Lu J-T, Pedemonte S, Bizzo B, Doyle S, Andriole K, Michalski M, et al. DeepSPINE: Automated Lumbar Vertebral Segmentation, Disc-level Designation, and Spinal Stenosis Grading Using Deep Learning2018.
42. Natalia F, Meidia H, Afriliana N, Young JC, Yunus RE, Al-Jumaily M, et al. Automated measurement of anteroposterior diameter and foraminal widths in MRI images for lumbar spinal stenosis diagnosis. PLoS One. 2020;15(11):e0241309.
43. Pang C, Su Z, Lin L, Lin G, He J, Lu H, et al. Automated measurement of spine indices on axial MR images for lumbar spinal stenosis diagnosis using segmentation-guided regression network. Med Phys. 2023;50(1):104-16.
44. Su ZH, Liu J, Yang MS, Chen ZY, You K, Shen J, et al. Automatic Grading of Disc Herniation, Central Canal Stenosis and Nerve Roots Compression in Lumbar Magnetic Resonance Image Diagnosis. Front Endocrinol (Lausanne). 2022;13:890371.
45. Tang H, Pei X, Huang S, Li X, Liu C. Automatic Lumbar Spinal CT Image Segmentation With a Dual Densely Connected U-Net. IEEE Access. 2020;8:89228-38.
46. Won D, Lee HJ, Lee SJ, Park SH. Spinal Stenosis Grading in Magnetic Resonance Imaging Using Deep Convolutional Neural Networks. Spine (Phila Pa 1976). 2020;45(12):804-12.
47. Kafri ASA, Sudirman S, Hussain AJ, Al-Jumeily D, Fergus P, Natalia F, et al., editors. Segmentation of Lumbar Spine MRI Images for Stenosis Detection Using Patch-

Based Pixel Classification Neural Network. 2018 IEEE Congress on Evolutionary Computation (CEC); 2018 8-13 July 2018.
48. Huang C, Li SX, Caraballo C, Masoudi FA, Rumsfeld JS, Spertus JA, et al. Performance Metrics for the Comparative Analysis of Clinical Risk Prediction Models Employing Machine Learning. Circ Cardiovasc Qual Outcomes. 2021;14(10):e007526.
49. Stevens L, Mortazavi B, Deo R, Curtis L, Kao D. Recommendations for Reporting Machine Learning Analyses in Clinical Research. Circulation: Cardiovascular Quality and Outcomes. 2020;13.
50. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, et al. On evaluation metrics for medical applications of artificial intelligence. Scientific Reports. 2022;12(1):5979.
51. Sudirman S, Ala Al-Kafri, Natalia F, Meidia H, Afriliana N, Al-Rashdan W, et al. Lumbar Spine MRI Dataset. Version 2. Mendeley Data. 2019. doi:10.17632/k57fr854j2.2.
52. Sudirman S, Ala Al-Kafri, Natalia F, Meidia H, Afriliana N, Al-Rashdan W, et al. Label Image Ground Truth Data for Lumbar Spine MRI Dataset. Version 2. Mendeley Data. 2019. doi:10.17632/zbf6b4pttk.2.
53. Sudirman S, Ala Al-Kafri, Natalia F, Meidia H, Afriliana N, Al-Rashdan W, et al. Radiologists Notes for Lumbar Spine MRI Dataset. Version 2. Mendely Data. 2019. doi:10.17632/s6bgczr8s2.2.
54. SpineWeb: Collaborative Platform for Research on Spine Imaging and Image Analysis. Datasets. 2018. Available from: http://spineweb.digitalimaginggroup.ca/Index.php?n=Main.Datasets. [Accessed on: 29-January-2024].

11

CONCLUSION AND DISCUSSION

The findings of this thesis contribute to the knowledge of conservative management with transforaminal epidural steroid injections in patients with lumbar radiculopathy, as well as underline the potential of innovations with artificial intelligence in the diagnosis and prognosis of spinal stenosis, marking another step toward more patient-tailored treatment strategies.

BALANCING SURGERY AND CONSERVATIVE CARE IN LUMBAR RADICULAR PAIN: THE ROLE OF EPIDURAL STEROID INJECTIONS

Considerations for surgical intervention

The current national guidelines in the Netherlands recommend that patients presenting to their general practitioner with signs of lumbar radiculopathy undergo an initial period of six to eight weeks of conservative management, consisting of oral pain medication and physical therapy if necessary. If symptoms persist beyond this timeframe, the patient can be referred to the neurologist for further evaluation, including MRI examination. The neurologist may then refer the patient to an outpatient pain clinic or a neurosurgeon if a surgically treatable anatomical substrate is identified.

The majority of patients diagnosed with lumbar radiculopathy demonstrates a clinically relevant disc herniation (LDH) on MRI as confirmed by our retrospective cohort analysis (**Chapter 3**). Previous research has demonstrated that surgical intervention is not necessarily the optimal treatment strategy, as lumbar disc herniation is often a self-limiting condition. A randomized trial involving 283 patients with radicular symptoms of 6 to 12 weeks due to LDH found that early discectomy did not yield superior long-term outcomes compared to prolonged conservative management after one year of follow-up [1]. The mere benefit of surgical intervention was a more rapid resolution of symptoms, while overall treatment costs remained comparable between the two strategies [2]. In clinical practice, one in five patients with disc herniation opt for immediate surgery, whereas the majority – over 50% – choose to continue conservative therapy. Similarly, the natural history of and the relative effectiveness of surgical intervention for lumbar radicular pain due to spinal stenosis remain subjects of ongoing debate.

We demonstrated that a large proportion of patients without disc herniation have a degenerative stenosis as the underlying cause of their symptoms (**Chapter 3**). A 2016 Cochrane Review concluded that there is limited evidence supporting surgical treatment over conservative management in this patient

population, a conclusion largely attributed to the methodological heterogeneity and low quality of existing trials [3]. Yet, a consensus group from the World Federation of Neurosurgical Society Spine Committee in 2020 agreed on the recommendation of decompressive surgery, in particular for patients with moderate-to-severe symptoms, citing evidence that surgical intervention resulted in faster and significantly greater symptom relief than conservative treatment, especially in the long term [4]. However, this recommendation was based largely on an as-treated analysis from a single study that combined a randomized and an observational cohort, and was compromised by a high rate of treatment cross-over [5]. The consensus group acknowledged that conservatively managed patients also demonstrated symptom improvement, albeit at a slower rate. In conclusion, the recommendation for a surgical solution for lumbar radiculopathy due to stenosis has weak strength and there is a need for further high-quality research in this area.

The varying effectiveness of surgical intervention to remove the compressive lesion and the discrepancy between clinical symptoms and radiological findings has fuelled the debate on the true cause of radicular pain. At first, it was assumed that physical neural compromise due to disc herniation and stenosis was the sole cause of symptoms. However, multiple studies have demonstrated the lack of correlation between clinical symptoms and nerve root compression as this is a frequent radiological finding in asymptomatic patients [6-13]. Moreover, symptomatic patients may experience improvement in symptoms without intervention that alters the underlying pathology, or may not show any compressive pathology at all whilst suffering from severe radicular pain. Although a correlation was found between contact pressure and neurological impairment [14], it was suggested that something else instigated radicular pain. Inflammation was implicated in the pathophysiology of lumbar radicular pain rather than physical compression after histological evidence was found of inflamed posterior nerve roots during laminectomy [15]. This was supported by findings that chymopapain, an anti-inflammatory substance used for chemonucleolysis of disc herniation, could cause rapid relief of leg pain preceding any change in disc herniation size or degree of nerve root compression [16, 17]. Furthermore, several additional studies demonstrated the presence of high levels of pro-inflammatory enzymes (PLA_2) and cytokines (IL-1α, IL-1β, IL-6, IL-8, TNF-α and PGE2) in patients with radicular pain with different injured spinal tissues releasing different types of inflammatory factors [18-23]. Additionally, some evidence has suggested that immunological processes play a role in the instigation of radicular pain as antibody levels to cell components of the central and peripheral nervous system and markers of glial cell and nerve damage have been shown to be elevated in patients with disc herniation [24, 25]. Hence, it is now an accepted concept that

lumbar radicular pain is caused by a complex interplay of inflammation, immune factors and mechanical compression of the nerve root.

Conservative management: the promise and limitations of epidural steroid injections

Considering that, on average, conservative therapy may be as effective as surgery in the long term for patients with lumbar radicular pain, it represents a viable alternative – provided that the patient's symptoms can be adequately reduced. Patients presenting with lumbar radiculopathy are initially treated with oral analgesics, starting with paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs). However, a substantial proportion of patients are prescribed additional opioid analgesics. Our retrospective data indicate that 57% of patients presenting to the outpatient pain clinic were using some form of morphine-like analgesics (**Chapter 3**). In light of the ongoing opioid crisis and the well-documented risks of dependency and adverse effects associated with chronic opioid use, this reliance on opioid analgesia is undesirable. Moreover, despite this pharmacological treatment, many patients continued to seek further therapeutic options.

Epidural steroid injections (ESI), although not addressing the underlying pathology directly – they do not create space for the neural elements or remove the herniated disc – are intended to modulate the inflammatory and immunological responses assumed to contribute to radicular pain. Corticosteroids are primarily thought to inhibit PLA_2, an important enzyme in the pro-inflammatory cascade, resulting in the reduced excretion of hyperalgesic prostaglandins, thromboxanes, and leukotrienes, thereby alleviating symptoms and creating a more tolerable clinical condition for the patient [23, 26, 27]. Hence, when effective, ESIs constitute a valuable addition to the spectrum of non-surgical treatment modalities. Conversely, in cases where ESIs fail to provide sufficient symptom relief, patients may still proceed to surgical intervention.

The use of epidural steroid injections (ESIs) has seen increasing prevalence, particularly as the role of inflammation in the onset of radicular pain has become more widely recognized. The current literature, however, is not unequivocal regarding the effectiveness of ESI in patients with lumbar radiculopathy. Interestingly, despite the lack of evidence, this treatment has become increasingly popular over recent decades. Our meta-analysis comparing the effectiveness of ESI to placebo injections provides some evidence that ESI is superior as this group reported greater reductions in pain scores and improvement of physical functioning (**Chapter 2**). Though, this was limited to short-term follow-up and the minimally clinical important difference (MCID) was not met in every study.

The interpretation of these findings is further complicated by substantial heterogeneity across studies, as ESIs can be administered via different approaches (caudal, interlaminar, or transforaminal) and involve varying corticosteroid formulations and dosages.

Optimizing patient selection: who benefits most from transforaminal epidural injections?

The paucity of definitive evidence supporting the effectiveness of transforaminal epidural steroid injections (TEI) may, in part, stem from suboptimal patient selection. Hence, identifying prognostic factors to guide clinicians in determining which patients are most suitable candidates for TEI is essential. However, this objective is impeded by the dearth of comparable studies. In our meta-analysis, sensitivity analyses were performed to explore potential determinants of treatment response, suggesting that ESI may be more effective in patients with MRI-confirmed LDH compared to those diagnosed solely on clinical assessment (**Chapter 2**). However, due to the necessity of pooling all three epidural approaches and the limited number of studies included in each analysis, these findings should be interpreted with caution.

Contrastingly, our large retrospective cohort study demonstrated that the short-term effectiveness of TEI is independent of the presence of a disc herniation on MRI (**Chapter 3**). This is an interesting finding, as it challenges the prevailing but unsubstantiated assumption that TEI is primarily beneficial for patients with LDH. While LDH generally follows a self-limiting course with eventual spontaneous resolution, spinal stenosis caused by degenerative changes typically progresses over time. Given that the effects of corticosteroids after a single injection are expected to diminish after several weeks, one could argue that TEI may serve as a temporary bridge to spontaneous recovery in LDH patients, whereas in those with spinal stenosis, symptom relief may be transient, but symptoms will eventually recur. However, as mentioned before, spontaneous symptom resolution can also occur in patients with spinal stenosis, and the precise role of inflammation – as well as the extent to which TEI modulates this process – is uncertain. Interestingly, in our systematic review of prognostic factors for TEI, two studies suggested that patients with spinal stenosis, specifically those with foraminal stenosis, may benefit more than LDH patients (**Chapter 4**). A potential explanation is that degenerative stenosis causes more dynamic, movement-dependent compression and recurrent cycles of inflammation around the nerve root. These intermittent flare-ups of perineural inflammation and vascular congestion may be more amenable to corticosteroid treatment than the continuous inflammation and static compression from disc herniation. In addition, in foraminal stenosis, corticosteroids delivered to the

neuroforamen may be confined by the stenotic components, hence, increasing the local drug concentration around the affected nerve. In contrast, in central or lateral recess stenosis, the injectables have to diffuse reducing their localized anti-inflammatory effects. Furthermore, vascularization is higher centrally and in the lateral recess, where corticosteroids may be more rapidly absorbed into the circulation, whereas in foraminal stenosis, compression-induced venous congestion and impaired perineural circulation may promote corticosteroid accumulation, prolonging its anti-inflammatory effects. Finally, a shorter symptom duration was associated with improved outcomes, although this relationship was identified in only four out of ten studies, and, hence, cannot be considered an established prognostic variable. The inconsistency of these findings highlights the need for studies specifically designed to establish reliable predictors of TEI effectiveness.

Biomarkers of inflammation

Since administration of corticosteroids is aimed at attenuating neuroinflammation, it is reasonable to postulate that TEI would be most beneficial for patients in whom inflammatory mechanisms play a predominant role. Consequently, identifying reliable biomarkers indicative of increased inflammatory activity is imperative for optimizing patient selection for TEI.

Gadolinium-based contrast agents are frequently employed in MRI examinations to enhance visualization of neuroinflammatory processes. Theoretically, the presence of gadolinium enhancement in the nerve root could serve as a predictor of TEI effectiveness. This hypothesis is supported by a retrospective study reporting greater symptomatic improvement following TEI in patients with nerve root enhancement on MRI compared to those without contrast enhancement [28]. However, histological analyses of intervertebral disc samples obtained from surgical patients have not demonstrated significant differences in macrophage-mediated inflammation between contrast-enhanced and non-enhanced groups [29].

Another proposed imaging biomarker is the presence of Modic changes (MC) on MRI, which are postulated to represent vascular deficiency, inducing an inflammatory milieu in the vertebral endplates. MC can be classified into three types, although only two are commonly observed. Type 1 MC is thought to represent an acute inflammatory process characterized by bone marrow oedema, whereas Type 2 MC is associated with a more chronic inflammatory state involving fatty marrow changes [30-32]. Yet, there is evidence suggesting that type 1 can develop into type 2 over time. Studies have correlated the presence of MC with an inflammatory state of the intervertebral disc and altered macrophage

distribution, with a relative depletion of anti-inflammatory M2 macrophages compared to pro-inflammatory M1 macrophages [33]. Since corticosteroids have been demonstrated to drive macrophage differentiation towards the M2 type in a mice model with lung injury [34], TEI could hypothetically facilitate an anti-inflammatory shift. However, our prospective cohort study found no significant differences in patient outcomes following TEI between those with and without MC (**Chapter 7**). Nevertheless, these results must be interpreted with caution as there was a preponderance of MC type 2 (93.5%). Moreover, it is possible that in patients with MC type 2, recurrent inflammatory episodes occur – resulting in oedema (MC type 1) – but are not detectable as type 1, since fatty marrow transformation has already taken place, leading in a MC type 2 score. This suggests that MC classification may not reliably distinguish between active and past inflammation. We recommend that this correlation should be repeated in a larger study with a more balanced representation of MC types. Among MRI features, mean spinal nerve intensity from axial Dixon T2-weighted water-only images has been correlated with patient outcome after TEI, with higher nerve intensity being associated with greater symptomatic relief [35]. The authors suggested that axial T2-weighted fat-saturated MRI scans can detect perineural inflammatory and hyperaemic changes that may not be visible on routine T1- and T2-weighted imaging used for MC classification.

Beyond conventional MRI, positron emission tomography combined with magnetic resonance imaging (PET/MRI) has emerged as a promising modality for detecting inflammatory changes. This imaging technique makes use of radioligands targeting inflammatory markers to visualize neuroinflammation. A study by Albrecht et al. demonstrated elevated levels of the neuroinflammation marker 18-kDa translocator protein (TSPO) in the neuroforamen of patients with chronic radicular pain compared to healthy controls, with a spatial distribution corresponding to the side of leg pain [36]. Moreover, increased TSPO expression correlated with improved outcomes following ESI, although the number of patients receiving an injection in this study was limited. Similarly, Lutke Schipholt et al. used this imaging modality to quantify the degree of inflammation in the spinal cord and neuroforamen in patients that suffered from cervical radiculopathy, using unaffected neuroforamina within the same patient as control, and found elevated inflammation levels at the affected neuroforamina [37]. In a subsequent study, they confirmed these findings by demonstrating higher levels of inflammation in neuroforamina corresponding to symptomatic cervical radiculopathy when compared to asymptomatic controls [38]. This imaging technique holds promise for aiding clinical practice in identifying patients who may benefit from anti-inflammatory therapy. However, its current practical challenges and high costs may limit its widespread clinical application.

Serological biomarkers represent a more accessible and cost-effective alternative for assessing neuroinflammation. However, research in this domain remains limited in the context of lumbar radiculopathy. One study investigated the predictive value of high-sensitivity C-reactive protein (hsCRP) for TEI response in patients with chronic unilateral radiculopathy secondary to spinal stenosis [39]. Patients were stratified based on pre-treatment hsCRP levels, yet no significant differences in outcomes were observed, suggesting that this inflammatory marker may not be a reliable predictor of TEI effectiveness. A fundamental challenge in utilizing serological biomarkers lies in selecting an appropriate reference group for comparison. While TEI response could serve as a surrogate marker, this approach would not directly elucidate the underlying cellular and molecular mechanisms. More invasive methodologies, such as histological analysis of tissue samples obtained during disc surgery, could provide valuable insights, though they are inherently biased toward patients with severe, intractable radiculopathy who undergo surgical intervention.

An alternative, minimally invasive technique for assessing the inflammatory microenvironment surrounding the nerve root is epidural lavage. This method involves the injection and subsequent withdrawal of a small volume of fluid from the epidural space, allowing for the quantification of inflammatory cytokines and peptides [40]. Golish et al. employed epidural lavage to identify a complex of matrix protein degradation fragments associated with inflammatory cytokines as a potential biomarker for ESI response [41]. Although in a small study population, they demonstrated that patients with this complex demonstrated significantly greater functional improvement following ESI compared to those without it. Hence, this might hold promise as a reliable, minimally invasive technique to predict TEI effectiveness.

Advancing research and clinical implications

Although not unequivocally, the current literature provides evidence for trends suggesting that the effectiveness of TEI is largely independent of the underlying aetiology of lumbar radiculopathy, though it may be more effective in patients with a shorter duration of symptoms. Hence, we argue that in patients with a recent onset of a clinically distinct lumbar radiculopathy, the necessity for MRI examination before TEI may be redundant and this therapy can be offered at an earlier stage. This rationale has been the impetus for the TEIAS trial (Transforaminal Epidural Injection in Acute Sciatica) (**Chapter 5**), which aims to investigate the potential benefits of expedited TEI therapy. At present, the long waiting times for clinical consultations by secondary care specialists (i.e., neurologists, orthopaedics and/or neurosurgeons) and imaging assessments – often extending over several weeks or months – result in substantial delays in

treatment initiation. During this period, patients frequently remain impaired in daily functioning and are managed with opioid analgesics. In case of persistent symptoms, some patients may ultimately opt for surgery in pursuit of rapid symptom relief, although long-term patient outcomes do not differ significantly compared to conservative therapy. Hence, the inadequacy of current guidelines in addressing pain management in the acute phase highlights the need for alternative therapeutic strategies. If expedition of TEI proves effective, it could provide general practitioners with an additional therapeutical option to achieve improved symptom management and physical recovery. Simultaneously, this may reduce opioid consumption, work absenteeism and the likelihood of surgery, thereby saving substantial medical and socioeconomic costs.

The TEIAS trial remains ongoing due to the lower-than-anticipated patient enrolment and will be concluded upon reaching the target sample size. A similar study conducted in a regional hospital in the Netherlands was prematurely terminated due to insufficient patient inclusion [42]. This randomized controlled trial examined the effectiveness of TEI in patients with acute lumbar radicular pain (<8 weeks) secondary to MRI-confirmed disc herniation with nerve root impingement, comparing three treatment arms: transforaminal epidural injection with steroid and a local anaesthetic, an injection with saline and a local anaesthetic, and usual care. The protocol intended an enrolment of a total of 264 patients, but in four years the group was able to recruit 141 participants. The trial demonstrated a statistically significant but clinically marginal improvement in leg pain following TEI with steroids and a local anaesthetic compared to usual care (NRS overall mean difference −0.96 (95% CI −1.83 to -0.09; P = 0.030), with no superiority observed for other outcome measures. Moreover, no statistically significant differences were observed between injections with steroids and those with saline (NRS leg pain overall MD −0.28 (95% CI −1.50 to 0.95; P = 0.659). This may be explained by the possible washout effect of cytokines when injecting saline near the affected nerve root. Consequently, the authors concluded that TEI could not be considered effective. In addition, their cost-effectiveness analysis indicated that TEI was not a viable economic alternative [43]. However, due to the limited sample size – 53% of their intended inclusions – there was a lack of statistical power for one of the primary outcome measures. Furthermore, the results were possibly biased by the high loss to follow-up (up to 30%). Since other studies, as summarized in our meta-analysis (**Chapter 2**), were able to demonstrate a statistically significant superiority of steroids over epidural placebo, further research in this area is deemed appropriate.

Although data from the TEIAS trial will be used for subgroup analyses to identify prognostic determinants for responders and non-responders to TEI, the study is

not explicitly designed for that purpose. To address this gap, the POTEISS cohort study (Prediction of Transforaminal Epidural Injection Success in Sciatica) was initiated to develop a prediction model for patient outcomes after TEI (**Chapter 6**). This study systematically evaluates patient outcome after standardized TEI treatment in patients with MRI-confirmed LDH or spinal stenosis causing unilateral radicular symptoms. It was designed with an adequate sample size to facilitate robust prediction modelling, incorporating a comprehensive set of demographical, clinical and radiological variables. The ongoing POTEISS study holds the potential to enhance clinicians' ability to counsel patients on TEI and its expected therapeutical effects, representing an important step toward evidence-based decision-making and more individualized treatment strategies.

In addition to these two studies, future research endeavours should further investigate inflammatory biomarkers and their association with TEI outcomes. Additionally, MC subtypes and treatment response should be assessed in larger studies, ensuring adequate representation of MC type 1, and investigating the usability of other MRI sequences in detecting ongoing inflammation. Furthermore, histological and serological analyses, in the presence and absence of MC, should be correlated with TEI effectiveness. The ongoing EIMICOR cohort study aims to address these aspects and includes patients who have undergone TEI therapy [44]. In addition, epidural lavage represents a promising technique that warrants further validation, alongside advanced imaging modalities such as PET/MRI, which may offer novel insights into neuroinflammation and treatment response. Collectively, these efforts will refine our understanding of the therapeutic potential of TEI, optimize patient selection criteria, and ultimately improve clinical management strategies for lumbar radiculopathy.

THE DAWN OF ARTIFICIAL INTELLIGENCE IN LUMBAR SPINE RESEARCH

Automating diagnostics for spinal stenosis

As previously stated, a key objective in spine research is the advancement toward more personalized treatment strategies. To this end, it is essential to estimate the likelihood of treatment success given a patient's demographic, clinical and radiological profile. For radiological assessments to contribute meaningfully to predictive models, standardized and reliable grading systems are indispensable. Despite the development of comprehensive grading systems that evaluate LSS as a multifaceted degenerative condition, their clinical implementation has remained limited due to suboptimal inter-reader agreement and weak correlations with baseline symptoms and surgical outcomes.

The widely adopted Schizas classification for central canal stenosis (CSS), consisting of seven grades, has demonstrated variable, predominantly moderate, inter-reader agreement [45-49]. Guen et al. reported substantial to near perfect interobserver agreement for their grading system; however, external validation studies revealed only fair agreement, raising concerns about its generalizability [50, 51]. In contrast, the four-grade Miskin scale for CSS consistently demonstrates substantial inter-reader agreement, as confirmed in our validation study (**Chapter 9**), suggesting that it may serve as a more reliable and practical alternative. Moreover, our validation of the Miskin scale for lateral recess stenosis (LRS) also demonstrated substantial agreement, which, together with CCS, supports a structured framework for intra- and interdisciplinary communication regarding spinal stenosis. This is further reinforced by its development through multidisciplinary consultation between radiologists and spine surgeons.

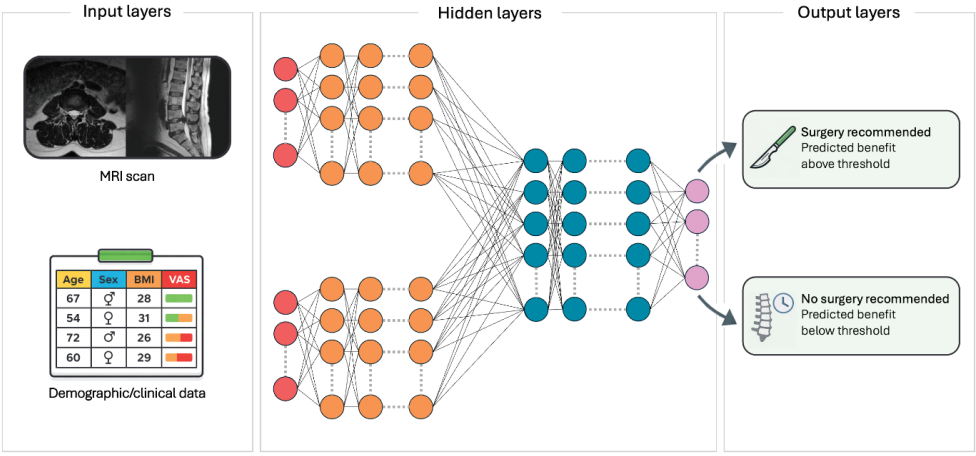
However, the need to eliminate inter-reader variability arises not only from the requirement for precise intra- and interdisciplinary communication but also from its potential contribution to the limited correlation between imaging findings and patient symptoms in previous studies [45-47, 52]. In our validation study of the Miskin scale, higher CCS grades were independently associated with greater improvement in leg pain, back pain and physical disability at 26 and 52 weeks after surgery, underscoring the clinical relevance and applicability of this grading system for preoperative counselling (**Chapter 9**). Nevertheless, clinically meaningful correlations were primarily observed across multiple severity categories, which may be partly attributable to residual inter-reader variability. Recent advances in artificial intelligence (AI), together with the exponential growth of data, offer potential avenues to address this limitation. In particular, AI-based techniques can be applied to automate the classification of radiological imaging, thereby eliminating inter-reader variability. Although machine learning (ML) algorithms have been in development since the 1950s, their recent integration into everyday applications accessible to the general public has sparked widespread interest. Together with the advent of advanced deep learning (DL) models, this has catalysed the use of AI in medical research. In recent years, the number of annual publications addressing segmentation and classification tasks for LSS using conventional ML or DL algorithms has increased substantially, with model performance nearing perfection (**Chapter 10**). However, while numerous novel algorithms have been developed, only few have been externally validated by other research groups using new datasets, limiting assessments of their robustness and generalizability. Additionally, the heterogeneous use of performance metrics presents a challenge in comparing model outcomes across studies. This underscores the need for a standardized set of evaluation metrics that should be consistently reported to ensure compa-

rability and reproducibility. Beyond eliminating inter-reader variability, AI-driven models have demonstrated the potential to significantly reduce workload and reporting time. Studies have shown that AI-assisted radiologists can reduce MRI reporting times by 56 seconds up to 14.1 minutes per study compared to non-AI-assisted radiologists [53, 54]. This efficiency gain represents a significant opportunity for clinicians to reallocate valuable time toward other critical tasks.

An innovative approach to spinal stenosis surgery outcome

Although existing grading scales for LSS severity incorporate multiple items, they may inadvertently omit critical information, thereby failing to correlate effectively with patient symptoms or treatment outcomes. Deep learning models in particular, hold great promise for revolutionizing such imaging analysis tasks. A single axial lumbar spine MRI slice contains 65,536 pixels on average assuming a 256x256 matrix – each pixel representing a specific grey-level intensity. This accumulation of data points results in a highly complex image, making it challenging for human observers to discern all relevant patterns. In contrast, DL algorithms are particularly well-suited for handling high-dimensional data as they can analyse vast numbers of data points simultaneously and identify intricate patterns that might be overlooked by the human eye. This capability was demonstrated in a study on the cervical spine, where a convolutional neural network (CNN) was used to predict clinical success following disc surgery from radiographs [55]. Intriguingly, the CNN's saliency maps revealed that the facet joints – a structure not typically emphasized by neurosurgeons when evaluating these patients for surgical candidacy – had a significant influence on its predictions. Given that current MRI-based degeneration scales for LSS provide limited guidance for surgical decision-making, applying a similar deep learning approach to lumbar spine MRI could uncover imaging features that may be relevant for prognosis but are not readily recognized by clinicians and enhance patient selection for surgery [56-58]. To explore this, we initiated a project to develop a model that uses T2-weighted MRI scans of the stenotic level of operated patients in combination with demographic and clinical data to predict whether a patient will benefit from spinal stenosis surgery. Model training utilized the patient-reported Zurich Claudication Questionnaire, which assesses pain, disability, and postoperative satisfaction at baseline and 26 weeks after surgery [59]. While such high-tech approaches are theoretically compelling, several practical challenges complicate their development. Data acquisition, for instance, is severely constrained by privacy regulations, limiting the size of available datasets. In addition, recruiting skilled engineers to commit to hospital-based research projects is difficult. Consequently, these constraints initially moderated the pace of progress, underscoring that developing a DL

model trained on multimodal data is a complex, time-intensive process that cannot be achieved quickly. Nevertheless, our first model, although not outperforming clinicians, yielded valuable insights that are guiding ongoing refinement and optimization.



Schematic view of a deep learning model for predicting surgical success for LSS patients

The road to AI-driven spinal care

Automating processes will be key in the near future to build a resilient and sustainable healthcare system. As healthcare systems worldwide face increasing pressure due to aging populations, rising costs, and workforce shortages, automation presents an opportunity to enhance efficiency, reduce administrative burdens, and improve patient outcomes. Reports have suggested that up to 35% of healthcare tasks could be automated, allowing medical professionals to dedicate more time to direct patient care [60]. While AI holds immense potential, its integration into clinical decision-making raises concerns about transparency, accountability, and patient trust. At the same time, translating AI models from research to routine neurosurgical practice requires overcoming significant barriers related to data availability, regulatory approval, and clinical validation. On the road to implementation and acceptance, addressing these challenges is crucial to ensuring that AI-driven advancements are both ethically sound and practically feasible.

Ethical frontiers

The integration of AI into medical decision-making presents a range of ethical challenges. A key concern is algorithmic bias, which arises when AI models reflect and potentially amplify existing disparities in healthcare data. Studies have shown that AI systems trained on unrepresentative datasets may perform less

accurately for underrepresented demographic groups, leading to inequitable patient outcomes [61, 62]. Addressing this bias requires diverse, high-quality training data and rigorous validation across different patient populations. Another critical issue is patient consent and transparency. Unlike traditional clinical decision-support tools, DL models function as ''black boxes,'' meaning their decision-making processes are often uninterpretable, even to their developers [63]. Although techniques, such as saliency maps or 'heat maps', have been developed to improve interpretability, they provide only limited insight into the underlying decision-making process and do not fully resolve the black-box nature of deep learning models. This lack of explainability complicates informed consent, as patients and clinicians may struggle to understand how an AI system arrives at its recommendations.

Accountability for AI-driven decisions further complicates matters. If an AI system makes an erroneous diagnosis or treatment recommendation, determining liability—whether it falls on the clinician, the hospital, or the developers of the AI system—remains a legal and ethical grey area [64]. Furthermore, the balance between automation and human oversight is crucial in ensuring AI augments rather than replaces clinical judgment. Over-reliance on AI could lead to deskilling of healthcare professionals, while excessive scepticism may hinder the adoption of beneficial algorithms [65]. Additionally, data privacy and security pose significant ethical challenges, as AI systems require vast amounts of sensitive patient data for training. Robust encryption, federated learning, and strict governance frameworks are necessary to prevent data misuse and unauthorized access [66].

From digital bench to bedside: challenges in AI implementation

The development and implementation of AI-driven algorithms into clinical practice face not only ethical concerns but also several practical challenges. A fundamental requirement for developing robust AI models is access to large, diverse, and high-quality datasets for effective training. However, the collection and sharing of such (anonymous) data is often impeded by privacy concerns, regulatory constraints, and the fragmented nature of healthcare data across institutions and regions. Legislation, such as the General Data Protection Regulation (GDPR) in Europe, adds complexity by imposing stringent data protection measures, including extensive data sharing agreements (DSA). While these measures are essential for safeguarding patient privacy, they strongly delay data exchange across departments, hospitals, and even countries, thereby slowing AI model development.

Once an AI model has been trained, extensive validation is required before it can be considered for clinical use. This requires rigorous assessment through clinical studies mimicking real-world practice to ensure it improves patient outcomes and does not introduce unintended harm or bias. The absence of standardized reporting metrics for AI performance further complicates this process, making it difficult to compare different models. Establishing a minimal set of outcome measures, as we recommended (**Chapter 10**), would facilitate more transparent comparisons and aid clinicians in selecting the most effective AI tools. Furthermore, the "black box" nature of AI models, while primarily an ethical issue, may also present practical challenges. The lack of explainability can hinder AI adoption, as both healthcare providers and patients may be reluctant to trust a system whose reasoning remains unclear. Furthermore, this lack of transparency may lead to hesitance among clinicians, who may prefer maintaining direct control over decision-making, unless they can interpret and justify the model's recommendations. The understanding and interpretation of AI recommendations hinges on the clinician's digital literacy, underscoring the need for integration of AI training into medical education to ensure clinicians are equipped with the necessary skills to effectively utilize AI tools, interpret their outputs, and understand their limitations [67].

Beyond model performance, successful implementation also depends on seamless integration into existing healthcare workflows. AI tools must complement, rather than disrupt, clinical processes, necessitating interoperability with electronic health record systems (HER) and intuitive user interfaces that align with clinicians' decision-making practices. Failure to achieve this alignment may lead to poor adoption rates, even if a model demonstrates clinical efficacy. Currently, this process is complex and time-consuming, requiring collaboration with each individual EHR software supplier to develop tailored integration solutions.

Once an AI model has been developed, validated, and optimized for clinical workflows, it must undergo regulatory approval before it can be deployed in practice. The European Union has introduced stringent standards that AI-driven technologies have to meet under the "Ethics Guidelines for Trustworthy AI", ensuring compliance with principles of transparency, accountability, and fairness, resulting in lengthy and resource-intensive approval processes [68].

Finally, financial constraints present a significant barrier to AI adoption in healthcare. The implementation of AI technologies requires substantial financial investments, yet many healthcare institutions face challenges to allocate sufficient funds for AI integration. Hospitals with limited financial resources often

find it difficult to attract and retain skilled computer engineers, whose expertise is essential for developing and implementing AI models in collaboration with clinicians. Additionally, commercially developed AI solutions can be prohibitively expensive, forcing hospitals to prioritize their investments and potentially limiting the breadth of AI adoption in clinical practice. To justify these investments, it is essential to demonstrate the cost-effectiveness of AI interventions through economic evaluations alongside clinical assessments. Hence, hospitals should be supported by governmental funding to facilitate broader adoption and equitable access to these technologies.

In the context of the aforementioned practical and ethical challenges, the deep learning model we aim to develop to predict surgical success for LSS represents an important first step toward AI-assisted patient selection in spinal surgery. Following model development, clinical translation will require further validation, refinement, and careful navigation of these challenges. Nonetheless, the insights generated thus far support the transformative potential of AI and its capacity to contribute to more precise and data-driven decision-making in spinal care.

REFERENCES

[1] Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, et al. Surgery versus prolonged conservative treatment for sciatica. N Engl J Med. 2007;356(22):2245–56.

[2] van den Hout WB, Peul WC, Koes BW, Brand R, Kievit J, Thomeer RT. Prolonged conservative care versus early surgery in patients with sciatica from lumbar disc herniation: cost utility analysis alongside a randomised controlled trial. Bmj. 2008;336(7657):1351–4.

[3] Zaina F, Tomkins-Lane C, Carragee E, Negrini S. Surgical versus non-surgical treatment for lumbar spinal stenosis. Cochrane Database Syst Rev. 2016;2016(1):Cd010264.

[4] Costa F, Alves OL, Anania CD, Zileli M, Fornari M. Decompressive Surgery for Lumbar Spinal Stenosis: WFNS Spine Committee Recommendations. World Neurosurg X. 2020;7:100076.

[5] Weinstein JN, Tosteson TD, Lurie JD, Tosteson A, Blood E, Herkowitz H, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis four-year results of the Spine Patient Outcomes Research Trial. Spine (Phila Pa 1976). 2010;35(14):1329–38.

[6] Karppinen J, Malmivaara A, Tervonen O, Pääkkö E, Kurunlahti M, Syrjälä P, et al. Severity of symptoms and signs in relation to magnetic resonance imaging findings among sciatic patients. Spine (Phila Pa 1976). 2001;26(7):E149–54.

[7] el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, Van der Kallen BF, van den Hout WB, Jacobs WC, et al. Magnetic resonance imaging in follow-up assessment of sciatica. N Engl J Med. 2013;368(11):999–1007.

[8] Athiviraham A, Yen D, Scott C, Soboleski D. Clinical correlation of radiological spinal stenosis after standardization for vertebral body size. Clin Radiol. 2007;62(8):776–80.

[9] Sirvanci M, Bhatia M, Ganiyusufoglu KA, Duran C, Tezer M, Ozturk C, et al. Degenerative lumbar spinal stenosis: correlation with Oswestry Disability Index and MR imaging. Eur Spine J. 2008;17(5):679–85.

[10] van Rijn JC, Klemetso N, Reitsma JB, Majoie CB, Hulsmans FJ, Peul WC, et al. Symptomatic and asymptomatic abnormalities in patients with lumbosacral radicular syndrome: Clinical examination compared with MRI. Clin Neurol Neurosurg. 2006;108(6):553–7.

[11] Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am. 1990;72(3):403–8.

[12] Boos N, Semmer N, Elfering A, Schade V, Gal I, Zanetti M, et al. Natural history of individuals with asymptomatic disc abnormalities in magnetic resonance imaging: predictors of low back pain-related medical consultation and work incapacity. Spine (Phila Pa 1976). 2000;25(12):1484–92.

[13] Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med. 1994;331(2):69–73.

[14] Takahashi K, Shima I, Porter RW. Nerve root pressure in lumbar disc herniation. Spine (Phila Pa 1976). 1999;24(19):2003–6.

[15] Lindahl O, Rexed B. Histologic changes in spinal nerve roots of operated cases of sciatica. Acta Orthop Scand. 1951;20(3):215–25.
[16] Kato F, Mimatsu K, Kawakami N, Iwata H, Miura T. Serial changes observed by magnetic resonance imaging in the intervertebral disc after chemonucleolysis. A consideration of the mechanism of chemonucleolysis. Spine (Phila Pa 1976). 1992;17(8):934–9.
[17] Sawin PD, Traynelis VC, Rich G, Smith BA, Maves TJ, Follett KA, et al. Chymopapain-induced reduction of proinflammatory phospholipase A2 activity and amelioration of neuropathic behavioral changes in an in vivo model of acute sciatica. J Neurosurg. 1997;86(6):998–1006.
[18] Saal JS, Franson RC, Dobrow R, Saal JA, White AH, Goldthwaite N. High levels of inflammatory phospholipase A2 activity in lumbar disc herniations. Spine (Phila Pa 1976). 1990;15(7):674–8.
[19] Franson RC, Saal JS, Saal JA. Human disc phospholipase A2 is inflammatory. Spine (Phila Pa 1976). 1992;17(6 Suppl):S129–32.
[20] Nygaard OP, Mellgren SI, Osterud B. The inflammatory properties of contained and noncontained lumbar disc herniation. Spine (Phila Pa 1976). 1997;22(21):2484–8.
[21] Takahashi H, Suguro T, Okazima Y, Motegi M, Okada Y, Kakiuchi T. Inflammatory cytokines in the herniated disc of the lumbar spine. Spine (Phila Pa 1976). 1996;21(2):218–24.
[22] Burke JG, Watson RW, McCormack D, Dowling FE, Walsh MG, Fitzpatrick JM. Intervertebral discs which cause low back pain secrete high levels of proinflammatory mediators. J Bone Joint Surg Br. 2002;84(2):196–201.
[23] McLain RF, Kapural L, Mekhail NA. Epidural steroid therapy for back and leg pain: mechanisms of action and efficacy. Spine J. 2005;5(2):191–201.
[24] Brisby H, Balagué F, Schafer D, Sheikhzadeh A, Lekman A, Nordin M, et al. Glycosphingolipid antibodies in serum in patients with sciatica. Spine (Phila Pa 1976). 2002;27(4):380–6.
[25] Brisby H, Olmarker K, Rosengren L, Cederlund CG, Rydevik B. Markers of nerve tissue injury in the cerebrospinal fluid in patients with lumbar disc herniation and sciatica. Spine (Phila Pa 1976). 1999;24(8):742–6.
[26] Collighan N, Gupta S. Epidural steroids. Continuing Education in Anaesthesia Critical Care & Pain. 2009;10(1):1–5.
[27] Baqai A, Bal R. The mechanism of action and side effects of epidural steroids. Techniques in Regional Anesthesia and Pain Management. 2009;13(4):205–11.
[28] Tak HJ, Jones R, Cho YW, Kim EH, Ahn SH. Clinical evaluation of transforaminal epidural steroid injection in patients with gadolinium enhancing spinal nerves associated with disc herniation. Pain Physician. 2015;18(2):E177–85.
[29] Djuric N, Yang X, Barzouhi AE, Ostelo R, Duinen SGV, Nijeholt G, et al. Gadolinium Enhancement Is Not Associated With Disc Inflammation in Patients With Sciatica. Spine (Phila Pa 1976). 2019;44(12):E742–e8.
[30] Dudli S, Sing DC, Hu SS, Berven SH, Burch S, Deviren V, et al. ISSLS PRIZE IN BASIC SCIENCE 2017: Intervertebral disc/bone marrow cross-talk with Modic changes. Eur Spine J. 2017;26(5):1362–73.
[31] Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disk disease. Radiology. 1988;168(1):177–86.

[32] Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology. 1988;166(1 Pt 1):193–9.
[33] Djuric N, Lafeber GCM, Li W, van Duinen SG, Vleggeert-Lankamp CLA. Exploring macrophage differentiation and its relation to Modic changes in human herniated disc tissue. Brain Spine. 2022;2:101698.
[34] Tu GW, Shi Y, Zheng YJ, Ju MJ, He HY, Ma GG, et al. Glucocorticoid attenuates acute lung injury through induction of type 2 macrophage. J Transl Med. 2017;15(1):181.
[35] Kim YH, Kim HJ, Seo J, Chai JW, Song HG, Choi YH, et al. Spinal nerve signal intensity on Dixon T2-weighted water-only sequence: an important outcome predictor after lumbar transforaminal epidural injection. Eur Radiol. 2021;31(12):9459–67.
[36] Albrecht DS, Ahmed SU, Kettner NW, Borra RJH, Cohen-Adad J, Deng H, et al. Neuroinflammation of the spinal cord and nerve roots in chronic radicular pain patients. Pain. 2018;159(5):968–77.
[37] Lutke Schipholt IJ, Scholten-Peeters G, Koop MA, Lammerstma AA, Coppieters MW, van de Giessen E, et al. Quantification of neuroinflammation in spinal cord and neuroforamina in patients with painful cervical radiculopathy using [11C]-DPA713 PET-CT. Submitted for publication.
[38] Lutke Schipholt IJ, Koop MA, Coppieters MW, van de Giessen EM, Lammerstma AA, Ter Meulen BC, et al. Neuroinflammation at the Neuroforamina and Spinal Cord in Patients with Painful Cervical Radiculopathy and Pain-Free Participants: An [(11)C]DPA713 PET/CT Proof-of-Concept Study. J Clin Med. 2025;14(7).
[39] Park CH, Lee SH. Prognostic usefulness of high sensitivity C-reactive protein for transforaminal epidural steroid injection in patients with radicular pain. Pain Med. 2011;12(2):219–23.
[40] Scuderi GJ, Brusovanik GV, Anderson DG, Dunham CJ, Vaccaro AR, Demeo RF, et al. Cytokine assay of the epidural space lavage in patients with lumbar intervertebral disk herniation and radiculopathy. J Spinal Disord Tech. 2006;19(4):266–9.
[41] Golish SR, Hanna LS, Bowser RP, Montesano PX, Carragee EJ, Scuderi GJ. Outcome of lumbar epidural steroid injection is predicted by assay of a complex of fibronectin and aggrecan from epidural lavage. Spine (Phila Pa 1976). 2011;36(18):1464–9.
[42] Ter Meulen BC, van Dongen JM, Maas E, van de Vegt MH, Haumann J, Weinstein HC, et al. Effect of Transforaminal Epidural Corticosteroid Injections in Acute Sciatica: A Randomized Controlled Trial. Clin J Pain. 2023;39(12):654–62.
[43] Ter Meulen BC, Maas ET, van der Vegt R, Haumann J, Weinstein HC, Ostelo R, et al. Cost-effectiveness of Transforaminal epidural steroid injections for patients with ACUTE sciatica: a randomized controlled trial. BMC Musculoskelet Disord. 2024;25(1):247.
[44] Djuric N, Lafeber G, van Duinen SG, Bernards S, Peul WC, Vleggeert-Lankamp CLA. Study protocol: effect of infection, Modic and inflammation on clinical outcomes in surgery for radiculopathy (EIMICOR). BMC Neurol. 2021;21(1):379.
[45] Schizas C, Theumann N, Burn A, Tansey R, Wardlaw D, Smith FW, et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphology of the dural sac on magnetic resonance images. Spine (Phila Pa 1976). 2010;35(21):1919–24.
[46] Moojen WA, Schenck CD, Lycklama À Nijeholt GJ, Jacobs WCH, Van der Kallen BF, Arts MP, et al. Preoperative MRI in Patients With Intermittent Neuro-

genic Claudication: Relevance for Diagnosis and Prognosis. Spine (Phila Pa 1976). 2018;43(5):348–55.
[47] Minetama M, Kawakami M, Teraguchi M, Matsuo S, Enyo Y, Nakagawa M, et al. MRI grading of spinal stenosis is not associated with the severity of low back pain in patients with lumbar spinal stenosis. BMC Musculoskelet Disord. 2022;23(1):857.
[48] Yang J, Xiong Y, Hu Y, Huang M, Zhang L, Pu X, et al. The reliability, correlation with clinical symptoms and surgical outcomes of dural sac cross-sectional area, nerve root sedimentation sign and morphological grade for lumbar spinal stenosis. BMC Musculoskelet Disord. 2023;24(1):225.
[49] Banitalebi H, Espeland A, Anvar M, Hermansen E, Hellum C, Brox JI, et al. Reliability of preoperative MRI findings in patients with lumbar spinal stenosis. BMC Musculoskelet Disord. 2022;23(1):51.
[50] Guen YL, Joon WL, Hee SC, Kyoung-Jin O, Heung SK. A new grading system of lumbar central canal stenosis on MRI: an easy and reliable method. Skeletal Radiology. 2011;40(8):1033–9.
[51] Al-Tameemi HN, Al-Essawi S, Shukri M, Naji FK. Using Magnetic Resonance Myelography to Improve Interobserver Agreement in the Evaluation of Lumbar Spinal Canal Stenosis and Root Compression. Asian Spine J. 2017;11(2):198–203.
[52] Pili M, Tobing SDAL. Functional outcome analysis of lumbar canal stenosis patients post decompression and posterior stabilization with stenosis grading using magnetic resonance imaging. Journal of Physics: Conference Series. 2017;884(1):012146.
[53] Lim DSW, Makmur A, Zhu L, Zhang W, Cheng AJL, Sia DSY, et al. Improved Productivity Using Deep Learning-assisted Reporting for Lumbar Spine MRI. Radiology. 2022;305(1):160–6.
[54] Yang YXC, Yee SY, Tan TSE, Koh KKN, Goh AGW, Venugopal VK, et al. An artificial intelligence boost to MRI lumbar spine reporting. Eur J Radiol. 2024;179:111636.
[55] Goedmakers C, Pereboom L, de Leeuw den Bouter M, Remis R, Staring M, Vleggeert-Lankamp C. Deep Learning on Preoperative Radiographs for Clinical Success Prediction after Surgery for Cervical Degenerative Disease. Brain and Spine. 2023;3:101842.
[56] Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, et al. Deep Learning in Medical Imaging. Neurospine. 2019;16(4):657–68.
[57] Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ Digit Med. 2021;4(1):65.
[58] Liu X, Gao K, Liu B, Pan C, Liang K, Yan L, et al. Advances in Deep Learning-Based Medical Image Analysis. Health Data Sci. 2021;2021:8786793.
[59] Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine (Phila Pa 1976). 1996;21(7):796–803.
[60] Spatharou A, Hieronimus S, Jenkins J. Transforming healthcare with AI: The impact on the workforce and organizations. McKinsey & Company; 2020.
[61] Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53.
[62] Colacci M, Huang YQ, Postill G, Zhelnov P, Fennelly O, Verma A, et al. Sociodemographic bias in clinical machine learning models: a scoping review of algorithmic bias instances and mechanisms. J Clin Epidemiol. 2025;178:111606.

[63] Samek W, Wiegand T, Müller K-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. arXivorg. 2017.
[64] Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. Artificial Intelligence in Healthcare. 2020:295–336.
[65] Choudhury A, Chaudhry Z. Large Language Models and User Trust: Consequence of Self-Referential Learning Loop and the Deskilling of Health Care Professionals. J Med Internet Res. 2024;26:e56764.
[66] Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. npj Digital Medicine. 2020;3(1):119.
[67] Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing Artificial Intelligence Training in Medical Education. JMIR Med Educ. 2019;5(2):e16048.
[68] European Commission. Ethics Guidelines for Trustworthy AI. 2019.

SUMMARY

Lumbar radiculopathy and neurogenic claudication are among the most common clinical syndromes encountered in spine care. Patients with these clinical syndromes typically present with severe leg pain, back pain, and impaired physical functioning, resulting in reduced quality of life. Despite advances in imaging, interventional pain management, and spine surgery, clinical decision-making remains challenging. Radiological findings do not always correlate with symptoms, treatment effects vary widely between patients, and clinicians lack reliable tools to predict outcomes of conservative or surgical interventions. This thesis aims to optimize care for patients with lumbar radiculopathy and neurogenic claudication by addressing two key domains: (1) the effectiveness and prognostic factors of transforaminal epidural steroid injections (TEI) in lumbar radiculopathy due to disc herniation or degenerative stenosis, and (2) the diagnostic and prognostic evaluation of lumbar spinal stenosis (LSS), with a focus on standardized grading systems and the potential role of artificial intelligence (AI).

Part I: optimizing care for lumbar radiculopathy

Lumbar radiculopathy, often referred to as sciatica, is characterized by unilateral radiating leg pain resulting from dysfunction of a lumbosacral nerve root. Although lumbar disc herniation (LDH) or degenerative stenosis are frequently identified on magnetic resonance imaging (MRI) as underlying causes, symptoms may also arise in the absence of clear compressive pathology, reflecting the complex interplay between mechanical compression, inflammation, and immunological processes. The natural course of lumbar radiculopathy is often favourable, particularly when caused by LDH, but spontaneous symptom resolution may also occur in cases related to degenerative stenosis. Consequently, current Dutch guidelines recommend an initial period of conservative management. However, this "wait-and-see" approach may expose patients to prolonged pain, functional limitations, and reduced quality of life.

Epidural steroid injections (ESI) are a minimally invasive treatment aimed at reducing nerve root inflammation and alleviating pain in lumbar radiculopathy, thereby providing symptomatic relief that makes the waiting period until spontaneous recovery or surgery more tolerable. Transforaminal epidural injections (TEI) are the most common approach, although caudal and interlaminar techniques are also used. Nevertheless, uncertainty persists in the literature regarding their effectiveness, optimal timing, and the patient characteristics that predict benefit. Our systematic review and meta-analysis of randomized controlled trials compared ESI, including all three approaches, with placebo injections in

patients with lumbar radiculopathy and demonstrated superior pain relief and functional improvement at 6 weeks and 3 months, although the magnitude of the difference in clinical effect was modest. The treatment effect was larger for TEI and caudal injections than for the interlaminar approach for short-term pain relief. Reported complication rates were low.

The effectiveness of TEI was evaluated in a retrospective cohort of patients who were referred to the outpatient pain clinic due to complaints of lumbar radiculopathy. This study focused on the effectiveness of TEI on pain reduction and whether outcomes depended on the presence of LDH on MRI. Approximately 70% of the 486 patients who underwent TEI experienced at least some short-term pain reduction, and, contrary to the common assumption that imaging-confirmed disc herniation is a prerequisite for benefit, clinical outcomes were comparable in patients with LDH and those with alternative or no clear compressive pathology. This finding challenges the routine reliance on MRI as a decisive factor for TEI eligibility and suggests that clinical assessment may be sufficient to guide initial treatment decisions.

Although the majority of patients benefit from TEI in the short term, the patient characteristics associated with a favourable response remain unclear. Identifying prognostic factors is essential to enable more tailored use of TEI. In our systematic review of prognostic factors associated with TEI outcomes, numerous demographic, clinical, and radiological variables were studied. Shorter duration of symptoms and, interestingly, lumbar radiculopathy secondary to stenosis rather than LDH emerged as potential predictors of a favourable treatment response, although they were not identified as such consistently across studies. In contrast, many commonly cited imaging features showed inconsistent or weak associations with treatment response. Overall, the review highlights the absence of robust, validated predictors and underscores the need for large, well-designed prospective studies to identify patient subgroups that are most likely to benefit from TEI.

One understudied imaging biomarker is the presence of Modic changes (MC). These vertebral endplate alterations are thought to reflect an inflammatory environment: MC type I represents a more acute process, whereas MC type II is associated with a more chronic inflammatory state and is the most prevalent type. In a prospective cohort study, the relationship between MC type II and the effect of TEI in 88 patients with lumbar radiculopathy secondary to LDH was examined, but no correlation was demonstrated. This suggests that this type of endplate change should not affect expectations regarding clinical outcomes

after TEI, further emphasizing the limited predictive value of isolated MRI features.

Considering the potentially limited value of MRI in relation to TEI effectiveness, the TEIAS trial has been initiated to evaluate whether early TEI in patients with acute lumbar radiculopathy provides more rapid and effective symptom relief than usual conservative care. In a subset of patients, symptom burden remains inadequately controlled with standard conservative management. This multicentre randomized controlled trial investigates whether early administration of TEI, based on clinical assessment without routine pre-treatment MRI, leads to superior symptom control during the early phase of disease compared with continued usual conservative care.

Additionally, the POTEISS study was designed to develop a multivariable prediction model for TEI treatment success, integrating clinical, demographic, and radiological factors. The results of this large prospective cohort study will aid clinicians in stratifying patients with lumbar radiculopathy due to LDH or degenerative stenosis for treatment with TEI, thereby facilitating more tailored treatment strategies.

Part II: improving diagnosis and prognosis in lumbar spinal stenosis

Lumbar spinal stenosis (LSS) is a degenerative condition characterized by narrowing of the spinal canal, lateral recess, or neuroforamen, often leading to neurogenic claudication. It represents one of the most common indications for spine surgery in older adults, yet postoperative outcomes and patient satisfaction vary widely. A major obstacle in the management of LSS is the lack of a universally accepted, clinically meaningful grading system that reliably correlates imaging findings with symptoms and outcomes.

Numerous grading systems for LSS have been proposed that assess the different anatomical regions of the lumbar spine contributing to this disease. The majority of grading systems are based on MRI examination. However, most systems assess only one anatomically relevant region, are subject to considerable intra- and inter-reader variability, and lack correlation with clinical parameters. A novel grading system was introduced by Miskin et al., assessing central canal stenosis (CCS), lateral recess stenosis (LRS), foraminal stenosis (FS), and facet arthropathy (FA). Using an independent cohort of patients with neurogenic claudication due to LSS who underwent surgery, the reliability and correlation with clinical data were evaluated. Substantial inter-reader agreement was demonstrated for CCS and LRS, while agreement for FS and FA remained limited.

Importantly, higher grades of CCS were independently associated with greater postoperative improvement in pain and disability, whereas LRS and a combined grading construct showed weaker clinical correlations. These findings suggest that CCS remains the most clinically relevant imaging parameter and that the proposed grading systems for FS and FA require refinement or substitution before broader implementation.

Recognizing the inherent variability and time burden of manual MRI grading, artificial intelligence (AI) has substantial potential to enhance LSS diagnostics. AI is capable of processing high-volume data in a short amount of time, detecting patterns that are difficult for humans to discern, and eliminating inter-reader variability. This makes AI particularly valuable in medical imaging, where it can improve diagnostic precision. A systematic review of existing algorithms was conducted to provide an overview of conventional machine learning (ML) and deep learning (DL) approaches for automated segmentation and classification of LSS. DL models demonstrated excellent capability in both segmentation and classification tasks, outperforming conventional ML methods, and nearing human-level performance. Moreover, these DL models can perform classification tasks without a separate segmentation step. However, the review also identified important limitations, including heterogeneity in outcome metrics, limited use of external validation, and lack of standardized test datasets. These shortcomings currently hinder clinical translation, despite the clear technical potential of AI-based tools.

SAMENVATTING

Lumbale radiculopathie en neurogene claudicatio behoren tot de meest voorkomende klinische syndromen binnen de wervelkolomzorg. Patiënten met deze klinische syndromen presenteren zich doorgaans met ernstige beenpijn, rugpijn en een verminderde fysieke functionaliteit, wat leidt tot een verminderde kwaliteit van leven. Ondanks ontwikkelingen in beeldvorming, pijnbehandelingen en wervelkolomchirurgie blijft de klinische besluitvorming uitdagend. Radiologische bevindingen correleren niet altijd met de klachten van de patiënt, behandelingsresultaten variëren sterk tussen patiënten en clinici beschikken over onvoldoende betrouwbare instrumenten om de uitkomst van conservatieve of chirurgische behandelingen te voorspellen. Dit proefschrift heeft tot doel de zorg voor patiënten met lumbale radiculopathie of neurogene claudicatio te optimaliseren door twee kerngebieden te onderzoeken: (1) de effectiviteit en prognostische factoren van een transforaminale epidurale steroïdinjectie (TEI) bij lumbale radiculopathie veroorzaakt door een hernia of een degeneratieve stenose, en (2) de diagnostische en prognostische evaluatie van patiënten met een lumbale spinale stenose (LSS), met een focus op de standaardisering van graderingsschalen en de potentiële rol van kunstmatige intelligentie (AI).

Deel I: optimalisatie van zorg voor patiënten met lumbale radiculopathie

Lumbale radiculopathie, in de volksmond vaker bekend als ischias, wordt gekenmerkt door unilaterale radiculaire pijn als gevolg van de disfunctie van een lumbosacrale zenuwwortel. Hoewel een lumbale hernia (LDH) of degeneratieve stenose op MRI regelmatig als onderliggende oorzaak wordt geïdentificeerd, kunnen symptomen ook optreden in afwezigheid van duidelijke compressieve pathologie op beeldvorming. De gedachte hierover is dat naast mechanische compressie van de zenuwwortel, inflammatie en immunologische processen een belangrijke rol spelen in het ontstaan van radiculaire klachten. Het natuurlijke beloop van lumbale radiculopathie is vaak gunstig, met name wanneer deze wordt veroorzaakt door een hernia, maar ook bij een degeneratieve stenose kan spontane verbetering van de klachten optreden. Daarom schrijven de huidige Nederlandse richtlijnen een initiële periode van conservatieve behandeling voor. Echter, deze "afwachtende" benadering kan patiënten blootstellen aan langdurige pijn, functionele beperkingen en een verminderde kwaliteit van leven.

Epidurale steroïdinjecties (ESI) vormen een minimaal invasieve behandeling die gericht is op het reduceren van inflammatie rondom de zenuwwortel en het verminderen van pijn bij lumbale radiculopathie. Doordat de pijn afneemt wordt de periode tot spontaan herstel of een eventuele operatie beter te ver-

dragen voor de patiënt. Transforaminale epidurale injecties (TEI) vormen de meest gebruikte benadering, hoewel ook caudale en interlaminaire technieken worden toegepast. Desondanks blijft er in de literatuur onduidelijkheid over de effectiviteit, de optimale timing van behandeling en de patiëntkenmerken die een gunstige respons na behandeling voorspellen. In een systematische review en meta-analyse van gerandomiseerde gecontroleerde studies werden ESI, waarbij alle drie de benaderingen werden meegenomen, vergeleken met placebo-injecties bij patiënten met lumbale radiculopathie. Hieruit bleek dat ESI na 6 weken en 3 maanden resulteerden in een betere pijnvermindering en functionele verbetering in vergelijking met placebo-injecties, hoewel de mate van verschil in klinische uitkomsten beperkt was. Het behandeleffect was groter voor transforaminale en caudale injecties dan voor de interlaminaire benadering voor wat betreft pijnvermindering op de korte termijn. In de studies werden weinig complicaties gerapporteerd.

De effectiviteit van transforaminale injecties werd verder onderzocht in een retrospectieve cohortstudie. In deze studie werden patiënten geïncludeerd die vanwege klachten passend bij lumbale radiculopathie werden verwezen naar de Pijnpolikliniek. Deze studie richtte zich op de effectiviteit van TEI op pijnreductie en op de vraag of de uitkomst afhankelijk was van de aanwezigheid van een hernia op MRI. Ongeveer 70% van de 486 patiënten die een transforaminale injectie kregen ervoer ten minste enige vermindering van pijn op de korte termijn. In tegenstelling tot de veel voorkomende aanname dat een hernia op de MRI een voorwaarde is voor effectieve behandeling met TEI, waren de klinische uitkomsten vergelijkbaar tussen patiënten met een hernia en patiënten met alternatieve of geen duidelijke compressieve pathologie. Met deze bevinding kan het routinematig gebruik van MRI als doorslaggevende factor voor de indicatiestelling van TEI worden betwist en kan worden overwogen dat de klinische beoordeling mogelijk voldoende is om initiële behandelbeslissingen te ondersteunen.

Hoewel het merendeel van de patiënten op de korte termijn profiteert van TEI, blijven de patiëntkenmerken die geassocieerd zijn met een gunstige behandelrespons onduidelijk. Het identificeren van prognostische factoren en daarmee een betere patiëntenselectie is essentieel om een meer gerichte toepassing van TEI mogelijk te maken. In onze systematische review naar prognostische factoren voor het effect van TEI werden talrijke demografische, klinische en radiologische variabelen onderzocht. Een kortere duur van de symptomen en, opvallend genoeg, lumbale radiculopathie veroorzaakt door een degeneratieve stenose in plaats van een hernia, kwamen naar voren als mogelijke voorspellers van een gunstige behandelrespons, hoewel deze bevindingen niet consistent

in alle geïncludeerde studies werden bevestigd. Daarentegen werden voor veel vaak genoemde kenmerken op MRI slechts zwakke of inconsistente correlaties met behandelrespons gevonden. Over het geheel genomen toont deze review het ontbreken van robuuste en gevalideerde voorspellende patiëntenkenmerken aan en onderstreept zij de noodzaak van grote, goed opgezette prospectieve studies om patiëntsubgroepen te identificeren die het meest waarschijnlijk baat hebben bij TEI.

Een weinig onderzocht MRI-kenmerk is de aanwezigheid van Modic-veranderingen (MC). Deze veranderingen van de dek- en eindplaten van de wervels worden verondersteld een inflammatoire omgeving te weerspiegelen: MC type I representeert een meer acuut proces, terwijl MC type II geassocieerd is met een meer chronische inflammatoire toestand en het meest frequent voorkomt. In een prospectieve cohortstudie werd de relatie tussen MC type II en het effect van TEI onderzocht bij 88 patiënten met lumbale radiculopathie ten gevolge van een hernia. Er werd echter geen duidelijke correlatie aangetoond. Dit suggereert dat dit type eindplaatverandering geen invloed zou moeten hebben op de verwachtingen ten aanzien het behandeleffect na TEI en benadrukt opnieuw de beperkte voorspellende waarde van geïsoleerde MRI-kenmerken.

Gezien de mogelijk beperkte waarde van MRI in relatie tot de effectiviteit van TEI is de TEIAS-studie opgezet om te evalueren of een vroege behandeling met TEI bij patiënten met acute lumbale radiculopathie leidt tot snellere en effectievere symptoomverlichting dan gebruikelijke conservatieve zorg. Bij een deel van de patiënten met acute klachten van lumbale radiculopathie is de symptoomlast onvoldoende onder controle met standaard conservatieve therapie zoals orale analgetica en fysiotherapie. Deze multicenter gerandomiseerde gecontroleerde studie onderzoekt of vroege behandeling met TEI, gebaseerd op klinische beoordeling zonder routinematige MRI voorafgaand aan de behandeling, leidt tot betere symptoomcontrole in de acute fase in vergelijking met voortzetting van de gebruikelijke conservatieve behandeling.

Daarnaast is de POTEISS-studie opgezet met als doel een multivariabel predictiemodel te ontwikkelen voor het voorspellen van behandelsucces van TEI, waarbij klinische, demografische en radiologische factoren worden geïntegreerd. De resultaten van deze grote prospectieve cohortstudie zullen clinici ondersteunen bij het stratificeren van patiënten met lumbale radiculopathie ten gevolge van een hernia of een degeneratieve stenose voor behandeling met een transforaminale injectie, en daarmee bijdragen aan meer gepersonaliseerde behandelstrategieën.

Deel II: verbetering van de diagnose en prognose bij lumbale spinale stenose

Lumbale spinale stenose (LSS) is een degeneratieve aandoening die wordt gekenmerkt door vernauwing van het centrale wervelkanaal, de laterale recessus of het neuroforamen, wat vaak leidt tot neurogene claudicatio klachten. Het vormt een van de meest voorkomende indicaties voor wervelkolomchirurgie bij oudere volwassenen, terwijl postoperatieve uitkomsten en patiënttevredenheid sterk variëren. Een belangrijk obstakel in de behandeling van LSS is het ontbreken van een universeel geaccepteerd en klinisch betekenisvol graderingssysteem voor MRI-scans dat de MRI-bevindingen betrouwbaar correleert met symptomen en klinische uitkomsten.

Er zijn verschillende graderingssystemen voor LSS voorgesteld die de verschillende anatomische regio's van de lumbale wervelkolom beoordelen die bijdragen aan deze aandoening. Het merendeel van deze systemen is gebaseerd op MRI-onderzoek. De meeste systemen beoordelen echter slechts één anatomisch relevante regio, zijn onderhevig aan aanzienlijke variatie in intra- en interbeoordelaarsbetrouwbaarheid en vertonen een beperkte correlatie met klinische parameters. Een nieuw graderingssysteem is recentelijk geïntroduceerd door Miskin et al., waarbij centrale kanaalstenose (CCS), laterale recessusstenose (LRS), foraminale stenose (FS) en facetartropathie (FA) worden beoordeeld. Met behulp van een onafhankelijk cohort van patiënten met neurogene claudicatio ten gevolge van LSS die een operatie ondergingen, werden de betrouwbaarheid en de correlatie met klinische gegevens geëvalueerd van dit nieuwe graderingssysteem. De interbeoordelaarsbetrouwbaarheid voor CCS en LRS was substantieel, terwijl de overeenstemming tussen beoordelaars voor FS en FA beperkt bleef. Belangrijk is dat ernstigere maten van CCS onafhankelijk geassocieerd waren met een grotere postoperatieve verbetering in pijn en functionele beperkingen, terwijl LRS en een gecombineerde graderingsscore van CCS en LRS zwakkere klinische correlaties vertoonden. Deze bevindingen impliceren dat CCS de meest klinisch relevante beeldvormingsparameter is voor patiënten met neurogene claudicatio en dat de voorgestelde graderingssystemen voor FS en FA verdere verfijning of vervanging vereisen voordat zij breder kunnen worden toegepast.

Gezien de variabiliteit in beoordeling en de tijd die het kost om MRI-scans handmatig te graderen heeft kunstmatige intelligentie (AI) een enorm potentieel om de diagnostiek van LSS te verbeteren. AI is in staat grote hoeveelheden data in korte tijd te verwerken, patronen te detecteren die voor mensen moeilijk waarneembaar zijn en interbeoordelaarsvariatie te elimineren. Hierdoor is AI bijzonder waardevol binnen de medische beeldvorming, waar het de diagnost-

ische nauwkeurigheid kan verbeteren. Een systematische review van bestaande algoritmen werd uitgevoerd om een overzicht te geven van conventionele machine-learning (ML)- en deep-learning (DL)-modellen voor geautomatiseerde segmentatie en classificatie van LSS. De DL-modellen toonden uitstekende prestaties bij zowel segmentatie- als classificatietaken, presteerden beter dan conventionele ML-methoden en benaderden het niveau van menselijke beoordelaars. Bovendien kunnen deze DL-modellen classificatietaken uitvoeren zonder een afzonderlijke segmentatiestap. Echter, deze review toonde ook belangrijke beperkingen aan in de huidige literatuur rondom AI-modellen, waaronder de heterogeniteit in uitkomstmaten om de prestaties van de modellen te beschrijven, de beperkte externe validatie van algoritmes en het ontbreken van gestandaardiseerde test-datasets. Deze tekortkomingen belemmeren momenteel de klinische implementatie, ondanks het duidelijke technische potentieel van AI-gebaseerde toepassingen.

ACKNOWLEDGEMENTS

Prof. dr. Vleggeert-Lankamp, beste Carmen, ik wil je graag bedanken voor de kans die je mij hebt geboden om dit promotieonderzoek te doen en de vrijheid die ik daarin had om eigen ideeën uit te werken en klinische studies op te zetten. Jouw tomeloze energie werkt inspirerend en ik hoop de komende jaren onze samenwerking in het onderzoek voort te zetten.

Prof. dr. Staring, beste Marius, veel dank voor jouw begeleiding en kritische blik, met name op het gebied van AI en de medische wetenschap. Ik waardeer het zeer dat ik van jouw kennis en ervaring heb kunnen leren en hoop dat er in de toekomst verdere stappen gezet kunnen worden om de brug tussen neurochirurgie en AI te versterken.

Beste Diederick en Niek, dank dat jullie mijn paranimfen willen zijn. Diederick, het is altijd handig dat ik met jou, als broer en collega-arts, kan sparren over onze carrières in het ziekenhuis. Niek, als collega-onderzoeker van het eerste uur hebben we hetzelfde traject doorlopen, en ik vind het dan ook prachtig dat we dit jaar allebei zijn gestart met de opleiding tot neurochirurg.

Dr. Smith, dear Timothy, thank you for giving me the opportunity to conduct research at CNOC and for allowing me to experience firsthand how neurosurgery is practiced in Boston.

Beste collega's van het LUMC en het HMC, in het bijzonder Louis, Samuel, Judith, Inge, Niek, Fleur, Caroline en Anna-Boukje, dank voor alle gezelligheid en leuke momenten de afgelopen jaren, tijdens zowel onderzoeksperiodes als in de kliniek, maar vooral ook buiten het ziekenhuis.

Beste staf en collega's van het Radboud, mijn transfer van de Randstad naar Nijmegen is nog niet al te lang geleden, maar ik voel mij al sinds de eerste dag ontzettend thuis. Veel dank voor de geweldige werksfeer en ik kijk er naar uit om de komende jaren veel mooie momenten met jullie te beleven.

Dank aan alle medeauteurs, betrokken collega-onderzoekers, collega-artsen in het Spaarne Gasthuis, het Alrijne Ziekenhuis en het Maasstad Ziekenhuis, en uiteraard aan alle patiënten die hebben deelgenomen aan onze studies.

Lieve vrienden en familie, dank voor jullie onvoorwaardelijke steun en vooral voor de noodzakelijke afleiding. In het bijzonder Mam, Pap, Died en Mau: jullie hebben mij bij iedere carrièrekeuze gesteund, ongeacht de omstandigheden. Zonder zulke lieve ouders en broers was ik niet gekomen waar ik nu sta en had ik niet kunnen doen wat ik het liefst doe.

Tot slot, lieve Am, dankzij jou heb ik dit succes kunnen bereiken. Met jou kan ik al mijn pieken en dalen delen, in Nederland en daarbuiten. Jouw liefde en vertrouwen geven mij de ruimte om mijn doelen na te streven. Daarom ben ik dankbaar dat ik dit met je kan delen en zie ik het eigenlijk als ons succes.

CURRICULUM VITAE

Eduard Verheijen was born in The Hague, the Netherlands on October 24, 1993. In 2010, he obtained his high school diploma from Gymnasium Beekvliet in Sint-Michielsgestel. Subsequently, he started the pre-medical major at Elizabethtown College in Elizabethtown, Pennsylvania, on a tennis scholarship. After this early sabbatical, he started his studies in Life Science and Technology at Leiden University and Delft University of Technology. In 2015, he earned his bachelor's degree following the completion of his thesis on the synergistic interaction between two chemotherapeutic agents for colon carcinoma at the Department of Pathology. He then continued his education at medical school in Leiden while simultaneously working at the Department of Radiology. During his undergraduate studies, he participated in the Honours College program, attended various extracurricular courses, and began research at the Department of Neurosurgery. This eventually led to the start of his MD-PhD track under the supervision of prof. dr. C.L.A. Vleggeert-Lankamp. His research focused on the value of transforaminal epidural injections in patients with a lumbar disc herniation or stenosis, leading to the initiation of two prospective clinical studies. During this period, he also completed several courses on machine learning and deep learning and participated in a four-month project aimed at optimizing workflow in the Radiology Department through artificial intelligence (AI). After completing medical school, he travelled abroad in the summer of 2023 to conduct research at the Computational Neuroscience Outcomes Center (CNOC), a scientific institution at Brigham and Women's Hospital in Boston. During his six-month stay, he investigated the application of AI in lumbar spine research. His research has resulted in multiple presentations at national and international conferences, including EUROSPINE 2021, the 47th ISSLS Annual Meeting, the EANS 2021 Congress, GSC 2024, and the Dutch Spine Society Annual Meeting. From February 2024 to January 2025, he worked as a resident-not-in-training in the Department of Neurosurgery at Haaglanden Medical Center. He continued in this role from March 2025 to December 2025 in the Department of Neurosurgery, Radboud University Medical Center. In January 2026, he started the neurosurgical residency program in Nijmegen.

LIST OF PUBLICATIONS

1. **Verheijen E**, Munts AG, van Haagen O, de Vries D, Dekkers O, van den Hout W, Vleggeert-Lankamp C. *Transforaminal epidural injection versus continued conservative care in acute sciatica (TEIAS trial): study protocol for a randomized controlled trial*. BMC Neurol. 2019 Sep 3;19(1):216. doi: 10.1186/s12883-019-1445-9. PMID: 31481010; PMCID: PMC6719366.
2. **Verheijen EJA**, Munts AG, van Haagen OBHAM, de Vries D, Vleggeert-Lankamp CLA. *The Outcome of Epidural Injections in Lumbar Radiculopathy Is Not Dependent on the Presence of Disc Herniation on Magnetic Resonance Imaging: Assessment of Short-Term and Long-Term Efficacy*. World Neurosurg. 2021 Apr;148:e643-e649. doi: 10.1016/j.wneu.2021.01.051. Epub 2021 Jan 23. PMID: 33497827.
3. **Verheijen EJA**, Bonke CA, Amorij EMJ, Vleggeert-Lankamp CLA. *Epidural steroid compared to placebo injection in sciatica: a systematic review and meta-analysis*. Eur Spine J. 2021 Nov;30(11):3255-3264. doi: 10.1007/s00586-021-06854-9. Epub 2021 May 11. PMID: 33974132.
4. **Verheijen EJA**, van Haagen OBHAM, Bartels EC, van der Sloot K, van den Akker-van Marle ME, Steyerberg EW, Vleggeert-Lankamp CLA. *Prediction of transforaminal epidural injection success in sciatica (POTEISS): a protocol for the development of a multivariable prediction model for outcome after transforaminal epidural steroid injection in patients with lumbar radicular pain due to disc herniation or stenosis*. BMC Neurol. 2024 Aug 20;24(1):290. doi: 10.1186/s12883-024-03801-1. PMID: 39164613; PMCID: PMC11334316.
5. **Verheijen EJA**, Kapogiannis T, Munteh D, Chabros J, Staring M, Smith TR, Vleggeert-Lankamp CLA. *Artificial intelligence for segmentation and classification in lumbar spinal stenosis: an overview of current methods.* Eur Spine J. 2025 Mar;34(3):1146-1155. doi: 10.1007/s00586-025-08672-9. Epub 2025 Jan 30. Erratum in: Eur Spine J. 2025 May;34(5):2040. doi: 10.1007/s00586-025-08783-3. PMID: 39883162.
6. **Verheijen EJA**, van der Vlist NRE, Bartels EC, van Haagen OBHAM, Vleggeert-Lankamp CLA. *The Effect of a Transforaminal Epidural Injection in Patients with Lumbar Disc Herniation is not correlated with the Presence of Type II Modic Changes*. Brain Spine. 2025 Feb 24;5:104222. doi: 10.1016/j.bas.2025.104222. PMID: 40115868; PMCID: PMC11925183.
7. **Verheijen EJA**, van der Vlist NRE, Vleggeert-Lankamp CLA. *Factors associated with patient outcome after transforaminal epidural injection for lumbar disc herniation and stenosis: a systematic review of current literature*. Submitted.

8. **Verheijen EJA**, Lindberg R, Lycklama à Nijeholt GJ, van der Kallen BF, Vleggeert-Lankamp CLA. *External Validation of a Novel Comprehensive Grading System for Lumbar Spinal Stenosis*. Submitted.
9. Wesselink EO, **Verheijen EJA**, Coppieters MW, Elliott JM, Weber II KA, Moojen W, Vleggeert-Lankamp C, Djuric N, Pool-Goudzwaard A. *Lumbar multifidus intramuscular fat concentrations are associated with recovery following decompressive surgery for lumbar spinal stenosis. A longitudinal cohort study with 5-year follow-up*. Spine (Phila Pa 1976). 2026 Jan 1;51(1):25-33. doi: 10.1097/BRS.0000000000005408. Epub 2025 May 27. PMID: 40433984; PMCID: PMC12226431.
10. **Verheijen EJA**, Taal W, Nandoe Tewari RDS, Giessen MCA, Roshani H. *Isolated cerebellar metastasis from urothelial carcinoma: A case report of a rare phenomenon*. Brain Spine. 2023 Nov 29;3:102716. doi: 10.1016/j.bas.2023.102716. PMID: 38105804; PMCID: PMC10724209.