Journal of Orthopaedic Surgery and Research

RESEARCH

Open Access

Applicability of randomized controlled trial evidence on surgery for lumbar disc herniations to clinical reality: a comparison with the nationwide FinSpine registry

Ida K. Rantalaiho[1*], Katri I. Pernaa[1], Jukka M. Huttunen[2,3], Nikolai Klimko[2,3], Henri Salo[4] and Antti V. Malmivaara[5,6]

Abstract

Aims The evidence of effectiveness of operative treatment of lumbar disc herniation (LDH) is based on findings from RCTs published during the last couple of decades. The applicability of these findings to clinical practise has not previously been evaluated by comparison with nationwide registry data. Our purpose was to assess the clinical homogeneity of the randomised controlled trials, to evaluate the applicability of the findings to the clinical reality utilizing FinSpine registry data and to propose ways to improve the applicability in the future.

Methods A systematic literature search was undertaken to find the RCTs. The benchmarking method compatible with the CONSORT statement was used to document and compare the characteristics of the RCTs and FinSpine registry population.

Results Six RCTs comparing operative treatment of LDH to different methods of conservative treatment showed heterogeneity of patients' clinical characteristics in trial protocols and in actual experiments and a poor adherence to intervention groups. Patient groups were not representative of the catchment area in any of the RCTs. The completeness of documentation of clinically relevant characteristics was limited in all of the RCTs. Despite the deficiencies, the RCT results on operative treatment were comparable with FinSpine registry results.

Conclusions The clinical heterogeneity of the RCTs, non-representative patient populations, incomplete reporting of patient characteristics and poor adherence to treatment groups limit generalizability and applicability of the existing RCTs' results. Our findings mark a need for future pragmatic RCTs as well as clinical registry-based studies to improve the evidence for decision making in real-life settings.

*Correspondence:
Ida K. Rantalaiho
ida.rantalaiho@varha.fi
[1]Department of Orthopedics and Traumatology, Turku University Hospital and University of Turku, Turku, Finland
[2]Department of Neurosurgery, Kuopio University Hospital, Kuopio, Finland
[3]Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland
[4]Knowledge Brokers Department, Data and Analytics, Finnish Institute for Health and Welfare, Helsinki, Finland
[5]National Institute for Health and Welfare, Helsinki, Finland
[6]Orton Orthopaedic Hospital, Helsinki, Finland

Introduction

Randomised controlled trials (RCTs) provide the most reliable evidence of the effectiveness of operative treatment of lumbar disc herniation (LDH) compared to conservative treatment and several randomised controlled trials [1–6] and meta-analyses [7] have been published on this study question. Majority of the studies suggest that in a certain time frame operative treatment offers faster pain relief and better recovery from disability compared to conservative treatment [1–4]. Lumbar disc herniation is a manifestation of a general intervertebral disc degeneration and a lot has been published on the pathogenesis recently [8–11].

FinSpine is a nationwide computer-based surgical spine registry designed to collect data on patient characteristics, interventions and outcomes [12]. In FinSpine the only inclusion criteria is surgery for a spinal disorder and exclusion criteria do not exist. Finspine documents the real-life praxis where the decision to operate is made by each individual surgeon and consequently the indications have not been predetermined for the registry. The FinSpine is one of the registries maintained by the Finnish Institute for Health and Welfare, its contents have been created by the Finnish Spine Surgery Association, and a consensus of the items to be included has been reached within the orthopaedic and neurosurgical communities in Finland [12]. The data is in accordance with the ICHOM recommendations, and currently hospitals providing data for the registry undertake over 90% of the spinal operations in Finland. To be added, current tightened information security legislation on data security assures that all research data from Finland is securely produced.

The aim of this study was (1) to assess the clinical homogeneity of the patients, interventions and outcome measures in RCTs assessing effectiveness of surgery for LDH, (2) to evaluate the applicability of the findings from the RCTs in comparison with clinical reality utilizing the national FinSpine registry and (3) to propose ways to improve the applicability. For these aims we evaluated the existing *RCTs' study protocols* concerning patient characteristics, contents of interventions and outcome measurements; and the *RCTs in the actual experiment* in patient selection, characteristics of patients, adherence to interventions and outcome measurements, and analysed the trajectories of patients in the RCTs and in the FinSpine registry.

Methods

We intended to find all randomised controlled trials published in peer-reviewed journals that compared effectiveness of operative treatment on LDH to conservative, non-invasive treatment. We utilised the systematic literature search of Council for Choices in Health Care in Finland concerning lumbar disc herniation, their operative treatment and rehabilitation. The search was executed in Ovid MEDLINE, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts of Reviews of Effect (DARE), Health Technology Assessment (HTA) Database, NHS Economic Evaluation Database (NHS EED) and International Guideline Library (G-I-N) by an informatician in 2019. The search strategy for MEDLINE can be found in reference Fig. 1 and strategies for all the other databases in reference Attachment 1 [13]. A check-up search in Ovid MEDLINE was conducted by KP in January 2023 to include the latest publications on the subject. The original Finnish systematic review protocol and execution of the review concords with the requirements of the PRISMA statement. Although, the research questions of this present study focus on applicability of the RCT data.

Authors KP and IR independently screened all the records of the systematic search and selected eligible RCTs comparing operative treatment of LDH to any generally used non-invasive, conservative treatment methods in consensus. Only publications in English were considered.

The benchmarking method (BM), which is in accordance with the CONSORT statement, was used for assessment of generalizability of the findings of the selected RCTs [14–16]. All descriptive information was extracted by author IR. Also, the supplementary material of the RCTs was studied. The accuracy of the extracted data was rechecked twice and then checked by another author KP.

The FinSpine registry data of operatively treated LDH patients between years 2017 and 2022 from five university hospitals in Finland, and data from the RCTs were compared. The Finnish health care is publicly financed non-profit system that covers all citizens of Finland. All primary LDHs were included regardless of the spinal level. FinSpine does not include conservatively treated patients. The surgical indication in the FinSpine has been judged by each individual surgeon.

Results

Altogether six relevant RCTs were found [1–6]. Five of the studies were found in the literature search of the Council for Choices in Health Care in Finland [2–6] and sixth in the check-up search [1]. The inclusion and exclusion criteria as well as index and control interventions, and the outcome measurements of the RCTs are described in Table 1. FinSpine is also included in Table 1 which shows that no exclusion criteria exist. Intolerable pain is mentioned as an exclusion criterion only in the study by Österman et al. The corresponding information from the FinSpine is included in Table 1. The baseline

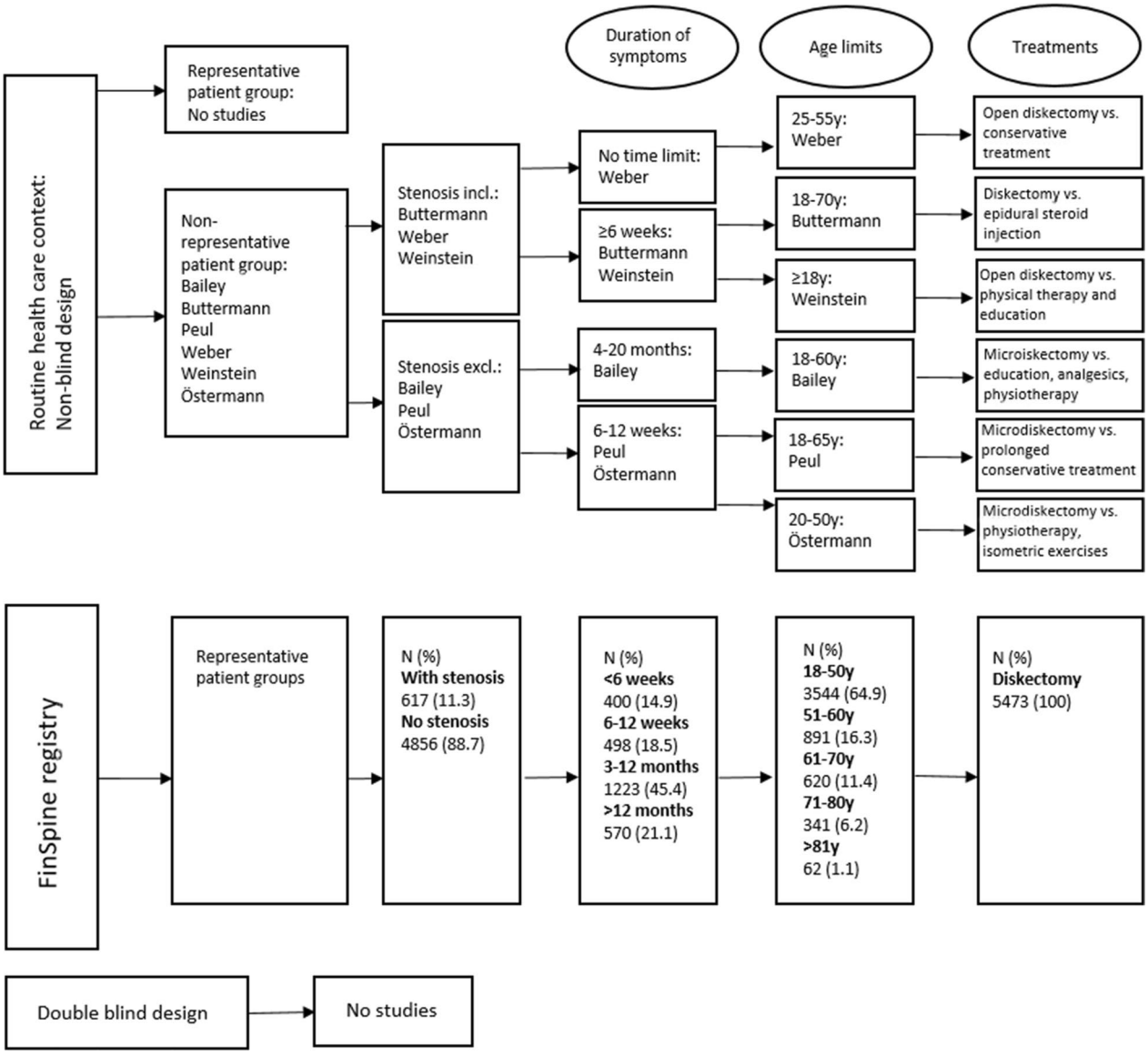

Fig. 1 Study question flowchart

characteristics of the patients in the RCTs and in the FinSpine are shown in Table 2.

The study question analysis flowchart showing trajectories of patients in each RCT and FinSpine is presented in Fig. 1. The trajectories of all the RCTs' differ from each other.

The completeness of documentation of the selection and characteristics of patients, adherence to interventions, categories of outcome measures, and use of statistical methods using the BM are shown in Table 3.

None of the studies reported patients' path before admission to the investigating hospitals. Österman et al. reported the catchment population of their primary investigating centre. Four trials reported the reasons for exclusion of patients before randomisation and the number of patients declining participation [1–3, 5]. All trials reported clinically important baseline characteristics and disorder specific clinical data, but only Peul et al. reported generic, and quality-of-life related functioning and Weinstein et al. was the only one to report some comorbid conditions. Both smoking and obesity was reported in three studies [1, 5, 6], one additional study reported smoking [2] and another reported obesity [3]. Any environmental factors such as work conditions, marital status, socioeconomic status, education or ethnicity was reported by four trials [1, 2, 4, 5].

All the studies reported adherence to the operative and conservative interventions, but only one reported the use of co-interventions during the follow-up (Tables 3 and 4) [6].

Table 1 Inclusion and exclusion criteria, index and control interventions, and planned outcome measures in RCTs on effectiveness of surgery on lumbar disc herniations and in the FinSpine registry

RCTs ↓	Inclusion criteria	Exclusion criteria	Index intervention	Control intervention	Outcome measures
Bailey et al. [1]	Age 18–60 y Unilateral radiculopathy of 4–20 mo Posterolateral disc herniation in the L4/5 or L5/S1 level disc on appropriate side and compression of the corresponding nerve root on MRI	Radiculopathy secondary to herniation of a foraminal or far lateral disc Spinal stenosis Deformity at the herniation level Previous lumbar surgery at the involved level Treatment of the current sciatica with epidural spinal injection Ongoing exercise-based physiotherapy	Microdiscectomy: an open or minimal-access approach	Nonsurgical treatment: education, oral analgesics, active physiotherapy Option for an epidural glucocorticoid injection 1–3 times	Primary: VAS leg pain at 6 mo Secondary: ODI, VAS back and leg pain, quality-of-life scores at 6 w, 3 mo, 6 mo, 1 y
Buttermann [2]	Lumbar disc herniation > 25% of the cross-sectional area of the spinal canal Duration of symptoms > 6 w	Age < 18 or > 70 y Pregnancy Cauda equina sydrome Pars defect at the level of the disc herniation Far-lateral disc herniation Multilevel symptomatic disc herniations Recurrent disc herniation	Discectomy	Epidural steroid injection 1–3 times one week apart	Motor strength scale 0–5, back and leg pain VAS, pain drawing, ODI, usage of pain medication, patient-assessed success of treatment at multiple timepoints
Peul et al. [3]	Age 18–65 y Radiologically confirmed disc herniation 6–12 w of incapacitating sciatica correlating MRI findings	Cauda equina syndrome Muscle paralysis or insufficient strength to move against gravity Occurrence of previous episode of similar symptoms within 12 mo Previous spine surgery Bony stenosis Spondylolisthesis Pregnancy Severe coexisting disease	Early sugery: microdiscectomy	Prolonged conservative treatment until 6 m: patient information, necessary pain medication, encouragement to return to daily activities, physiotherapy if necessary	Primary: Roland Disability Questionnaire, leg pain VAS and patient reported perceived recovery on 7-point Likert scale at multiple timepoints Secondary: SF-36, Sciatica frequency and bothersomeness index, health perception VAS at multiple timepoints
Weber [4]	Age 25–55 y Clinical symptoms and signs of a herniated lumbar disc, on the level of fifth lumbar and/or first sacral roots	Spondylolisthesis Previous spine operations	Open discectomy	Conservative treatment	Questionnaire at 3, 6 and 9 mo, 2 and 3 y Neurological examination and evaluation of psychosocial conditions at 1, 4 and 10 y
Weinstein et al. [5]	Age > 18 y Radicular pain for at least 6 w Evidence of nerve root irritation in clinical examination Disc herniation in vertebral imaging at the level and side corresponding to the clinical symptoms	Prior lumbar surgery Cauda equina syndrome Scoliosis greater than 15° Segmental instability Vertebral fractures Spine infection or tumor Inflammatory spondylarthropathy Pregnancy Comorbid conditions contraindicating surgery Inability/unwillingness to have surgery within 6 m	Open discectomy	Nonoperative treatment ("usual care"): including at least active physical therapy, education/counseling with home exercise instructions, NSAIDs if tolerated	Primary: changes from baseline for the Medical Outcomes Study 36-item Short-Form Health Survey bodily pain and physical function scales and ODI (MODEMS version) at 6 w, 3 mo, 6 mo, 1 y and 2 y Secondary: Sciatica frequency and bothersomeness indices, satisfaction with symptoms, self-reported improvement, employment status

Table 1 (continued)

RCTs ↓	Inclusion criteria	Exclusion criteria	Index intervention	Control intervention	Outcome measures
Österman et al. [6]	Age 20–50 y Below knee radicular pain of 6–12 w A CT finding of intervertebral disc extrusion or sequester At least one specific physical finding	Previous back surgery Spondylolisthesis Symptomatic spinal stenosis Over 3 m continuous sick leave due to back pain or leg pain preceding randomization Condition confounding evaluation of treatment outcomes (vascular claudication, symptomatic osteoarthritis, previous major trauma, diabetic polyneuropathy) A contraindication to conservative treatment (cauda equina syndrome, progressive neurologic deficit, intolerable pain)	Microdiscectomy within 2 weeks of randomization	Nonsurgical group: initially similar physiotherapeutic instructions and continued with isometric exercises	Primary: leg pain VAS at multiple timepoints Secondary: back pain VAS, work ability VAS, health-related quality of life 15D, ODI, satisfaction with treatment, global assessment of healing
FinSpine	Age ≥ 16 y MRI* confirmed primary disc herniation in the lumbar spine regardless of the level[a]	None	Discectomy		Leg and back pain VAS, ODI, satisfaction with treatment at 3 mo, 1, 2, 5 and 10 y

RCT randomized controlled trial, *y* years, *mo* months, *w* weeks, *MRI* magnetic resonance imaging, *VAS* visual analogue scale, *ODI* Oswestry disability index, *y* year, *MODEMS* modified version by American Academy of Orthopaedic Surgeons

*In the case of magnetic resonance imaging being contraindicated, a CT might have been used

[a]Only primary disc herniations included in this study although the FinSpine registry also includes data of recurrent herniations

Table 2 Baseline characteristics of the patients in the operative groups of the RCTs on effectiveness of surgery for lumbar disc herniations and in the FinSpine registry

RCT/hospital	Bailey et al. [1]	Buttermann (2004)	Peul et al. [3]	Weber [4]	Weinstein et al. [5]	Österman et al. [6]		FinSpine
Patients, n	64	50	141	60	232	28		5473
Female, n (%)	27 (42)	nr	52 (37)	28 (47)	101 (44)	13 (46)		2531 (46.2)
Age, mean (SD), y	38 (8.3)	40 (10)	41.7 (9.9)	40	41.7 (11.8)	37 (7)		46 (15.1)
BMI, mean (SD)	27.1 (5.6)	nr	25.9 (4.1)	nr	27.8 (5.6)	24 (4)		27.8 (4.9)
Current smoker, n (%)	22 (34.4)	18 (36)		nr	47 (20)	14 (50)		622*
Employment status, n (%)								
Employed	44 (68.8)	36 (20)	nr	nr	142 (61)	26 (93)		1552 (28.4)
Not employed	13 (20.3)	1 (2)	nr	nr	0 (0)	nr		212 (3.9)
Retired	0 (0)	0 (0)	nr	nr	0 (0)	nr		358 (6.5)
Disabled	5 (7.8)	13 (26)	nr	nr	27 (16)	nr		562 (10.3)
Other	2 (3.1)	0 (0)	nr	nr	63 (27)	nr		
MD	na	na	nr	na	na	na		2789 (51.0)
Comorbid conditions, n (%)	nr	nr	nr	nr	181 (78)**	nr		88*
Pain duration, mean (SD), wk	7.3 (2.5) mo[a]	3.8 (2.9) mo[a]	9.43 (2.37)[b]	nr	nr	11 (4.57)[b]	< 6 wk, n (%)	400 (7.3)
							6–12 wk, n (%)	498 (9.1)
							3–12 mo, n (%)	1223 (22.3)
							> 12 mo, n (%)	570 (10.4)
							md, n (%)	2782 (50.8)

RCT randomized controlled trial, *no.* number, *y* years, *BMI* body mass index, *md* missing data, *na* not applicable, *wk* weeks, *mo* months

*Percentages not announced due to missing data

**Any announced conditions, such as depression, joint problems, Parkinson's disease, heart or lung related problems, rheumatoid arthritis, stroke, diabetes etc.

[a]Bailey and Buttermann report duration of"symptoms" in months, duration of leg or back pain is not specified

[b]Peul and Österman report duration of leg pain

Table 3 Appropriateness of reporting of items enabling assessment of applicability of evidence of the RCTs on effectiveness of surgery for lumbar disc herniations and in Finnish university hospitals in the FinSpine registry

RCTs	Bailey et al. [1]	Buttermann (2004)	Peul et al. [3]	Weber [4]	Weinstein et al. [5]	Österman et al. [6]	FinSpine
Study characteristics ↓							
1. Selection of patients; health care system features							
1.1 Description of patients' path before assessment of eligibility	nr	nr	nr*	nr	nr	nr	nr
1.2 Reporting of reasons for exclusions before randomization	y	y	y	nr	y	nr	nr
1.3 Percentage of eligible patients declining participation documented	y	y	y	nr	y	nr	nr
1.4 Description of characteristics of all the health care settings where the data was collected	nr	nr	nr	nr	nr	nr	nr
2. Baseline characteristics of patients							
2.1 Demographic- and disorder-specific clinical data	y	y	y	y	y	y	y
2.2 Functioning (disease specific, generic, health-related quality of life) (at least 1; at least 2; all three items reported)	y	y	y	nr	y	y	y
	y	nr	y	nr	y	y	y
	nr	nr	y	nr	nr	nr	y
2.3 Comorbidity, at least two comorbid conditions reported or a comorbidity index	nr	nr	nr	nr	y	nr	y
2.4 Behavioral factors (smoking, alcohol/substance consumption, or exercise reported); and obesity	y;y	y;nr	nr;y	nr	y;y	y;y	y;y
2.5 Environmental factors (work or living conditions; marital status)	y	y	nr	nr	y	y	y
2.6 Potential inequity (socioeconomic status, education, deprivation, ethnicity)	y	nr	nr	y	y	y	nr
3. Interventions							
3.1 Content of the index intervention	y	y	y	y	y	y	y
3.2 Content or the control intervention	y	y	y	y	y	y	na
3.3 Staff competence	y	y	nr	nr	nr	y	y
3.4 Healthcare system features (e.g. resources, clinical paths)	nr	y	nr	nr	nr	nr	nr
3.5 Adherence to index intervention	y	y	y	y	y	y	na
3.6 Adherence to control intervention	y	y	y	y	y	y	na
3.7 Use of other healthcare services	nr	nr	nr	nr	nr	y	nr
4. Outcome data							
4.1 Primary and secondary outcomes	y	y	y	y	y	y	y
4.2 Percentage of and reasons for dropping out of follow-up	y;nr	y;nr	y;nr	nr	y;nr	y;nr	y;nr
5. Statistical analysis							
5.1 Description of power calculations	y	nr	y	nr	y	y	na
5.2 Description and appropriateness of all primary and secondary statistical analyses	y	y	y	y	y	y	na

y yes, *nr* not reported, *RCT* randomized controlled trial, *na* not applicable

*Described in a previously published protocol article

All studies reported the shares of crossing over from conservative groups to operative groups, as well as the reasons for crossing over (Table 4). Peul et al. and Österman et al. reported pre-specified criteria for crossing over to operative intervention. The proportions of patients crossing over to surgery varied between 26 and 54%; and crossing over to conservative varied between 0 and 44%. Buttermann had highest cross-over to surgery 54% and Weinstein et al. had the highest crossovers to both directions: 40% to surgery and 44% to conservative treatment. Two of the trials were conducted by single specialists [2, 4], while several specialists were included in the other trials [1, 3, 5, 6].

All trials reported their planned outcome measures and four reported primary and secondary outcomes separately (Table 1) [1, 3, 5, 6]. Five trials reported percentages of dropouts [1–3, 5, 6], but none reported reasons for dropping out. Two trials did not describe power calculations [2, 4] whereas four trials described power calculations appropriately [1, 3, 5, 6]. All trials used appropriate statistical analyses (Table 3).

In the FinSpine registry data 5473 patients were operatively treated for LDH (Table 2). Overall, the percentage of respondents for baseline information was 49.9 and the rate of respondents varied from 29.0 to 61.2% between the University Hospitals. The mean age of all the FinSpine registry patients was 46 (SD 15.1) years whereas Weinstein et al. and Peul et al. had the highest mean ages of the RCTs; 41.7 (SD 11.8) and 41.7 (SD 9.9) years, respectively, and Österman et al. had the lowest mean age of 37 (SD 7) years. The sex distribution was comparable in almost all the RCTs and in the FinSpine registry. In the FinSpine

Table 4 Crossovers and 1 year outcomes in the RCTs on effectiveness of surgery for lumbar disc herniations and in the FinSpine registry

RCT/hospital ↓	Cross overs, n (%)		Leg pain VAS*, mean (SD)		ODI[a], mean (SD)	
	Op → nonop	Nonop → op	Operative group	Nonoperative group	Operative group	Nonoperative group
Bailey et al. [1]	8 (13)	22 (34)	26 (4)	47 (4)	22.9 (2.3)	34.7 (2.4)
Buttermann (2004)	0 (0)	27 (54)	nr[b]	nr[b]	nr[b]	nr[b]
Peul et al. [3]	16 (11)	55 (39)	11.0 (1.9)	11.0 (1.9)	nr[c]	nr[c]
Weber [4]	1 (1.7)	17 (26)	nr	nr	nr	nr
Weinstein et al. [5]	107 (44)	103 (40)	nr[d]	nr[d]	16.9 (1.7)	18.9 (1.6)
Österman et al. [6]	0 (0)	11 (39)	6 (11)	9 (19)	10 (13)	11 (14)
FinSpine	na	na	25 (27)	na	16.9 (16.6)	na

RCT randomized controlled trial, *VAS* visual analogue scale range 0–100, *ODI* Oswestry disability index range 0–100, *nr* not reported, *na* not applicable

*Scores for VAS range from 0 to 100, with higher scores indicating more pain. In Baileys work values were transformed from 0 to 10 by multiplying with 10 for comparability

[a]Scores on the ODI range from 0 to 100, with higher scores indicating worse disability and pain

[b]Results reported only in bar charts, exact numbers not available

[c]Roland disability index reported: operative group 3.3 ± 0.5, nonoperative group 3.7 ± 0.5

[d]Sciatica frequency and bothersomeness indices reported

Table 5 Baseline values of leg pain (VAS) and ODI in RCTs on effectiveness of surgery for lumbar disc herniations and in the FinSpine registry

RCT/hospital ↓	Leg pain VAS*, mean (SD)		ODI[a], mean (SD)	
	Operative group	Nonoperative group	Operative group	Nonoperative group
Bailey et al. [1]	77 (20)	80 (18)	49.7 (15.8)	50.2 (15.9)
Buttermann (2004)	nr[b]	nr[b]	nr[b]	nr[b]
Peul et al. [3]	67.2 (27.7)	64.4 (21.2)	nr[c]	nr[c]
Weber [4]	nr	nr	nr	nr
Weinstein et al. [5]	nr[d]	nr[d]	47.5 (21.4)	46.3 (20.6)
Österman et al. [6]	61 (20)	57 (21)	39 (15)	39 (14)
FinSpine	69 (25)	na	46.8 (18.5)	na

RCT randomized controlled trial, *VAS* visual analogue scale range 0–100, *ODI* Oswestry disability index range 0–100, *na* not reported, *na* not applicable

*Scores for VAS range from 0 to 100, with higher scores indicating more pain. In Baileys work vaules were transformed from 0 to 10 by multiplying with 10 for comparability

[a]Scores on the ODI range from 0 to 100, with higher scores indicating worse disability and pain

[b]Results reported only in bar charts, exact numbers not available

[c]Roland disability index reported: operative group 16.5 ± 4.4, nonoperative group 16.3 ± 3.9

[d]Sciatica frequency and bothersomeness indices reported

registry data the pain duration is reported in four categories: less than 6 weeks, 6–12 weeks, 3–12 months and over a year. Of the FinSpine patients 11% were current smokers, whereas smoking varied from 20 to 50% in the RCTs. In the FinSpine the overall response rate for ODI at one 1 year was 41% and for VAS 38%.

Comparison of the baseline values of visual analogue scale (VAS) for leg pain and Oswestry Disability Index (ODI) between the selected RCTs and FinSpine registry is shown in Table 5.

Comparison of VAS for leg pain and ODI after interventions at 1 year is shown in Table 4. The mean VAS and ODI values at baseline and at 1 year follow-up are illustrated in Figs. 2 and 3.

Some studies did not report VAS for leg pain: Buttermann et al. only reported bar charts and not exact number, Weber did not report degree of leg pain at all, and Weinstein et al. used sciatica indices. ODI was reported in exact numbers by three studies: Bailey, Weinstein and Österman, while Buttermann again reported only bar charts. Weber did not report any patient reported outcome measure, and Peul et al. used Roland Disability Questionnaire score.

Discussion

To our knowledge, this is the first study to compare the characteristics of RCTs on effectiveness of LDH surgery to a nation-wide registry population of LDH surgery patients and the first study to assess the completeness of documentation of the RCTs.

Our study demonstrates heterogeneity in inclusion and exclusion criteria, actual study populations, intervention comparisons and outcome measurements between the RCTs. We found that the due to general deficiency in documentation and reporting of certain characteristics the applicability and generalizability of the results of the

Fig. 2 The mean ODI values in lumbar disc herniation patients at baseline and at 1-year follow-up after discectomy in previously published RCTs and FinSpine

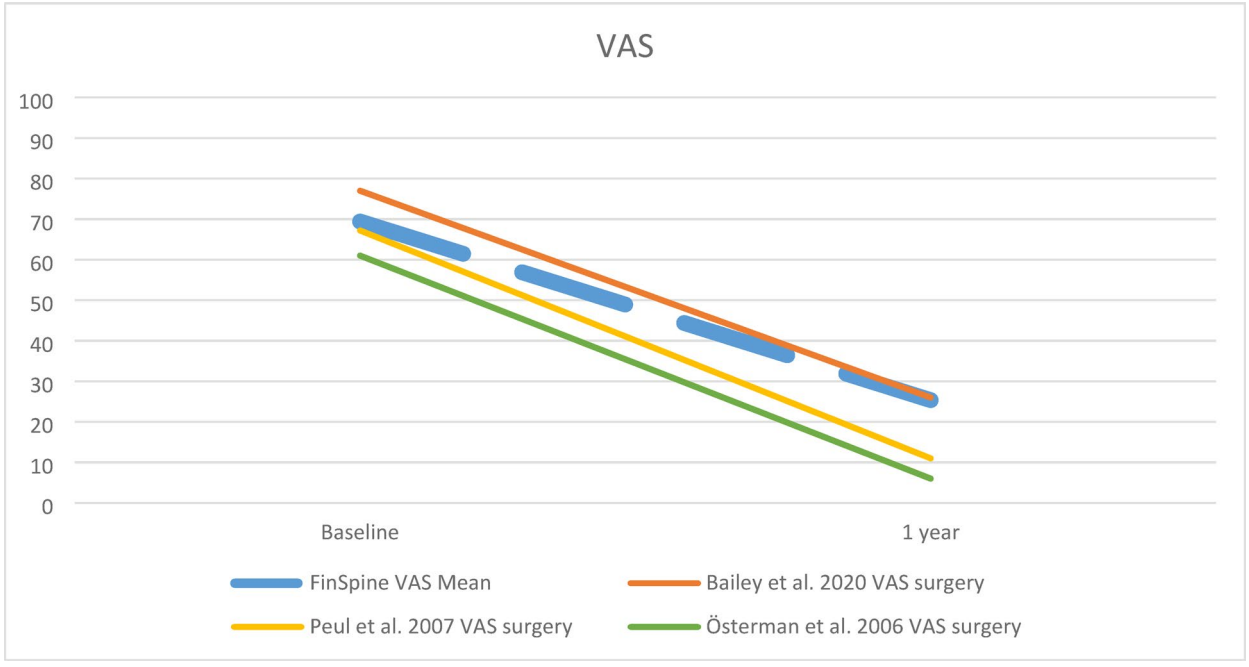

Fig. 3 The mean VAS leg pain values in lumbar disc herniation patients at baseline and at 1-year follow-up after discectomy in previously published RCTs and FinSpine

RCTs remains limited. If documentation and reporting were more precise in any future RCTs the applicability would be considerably improved.

We identified six relevant RCTs in the systematic literature search. Overall, the results in the RCTs showed no significant differences in recovery at 1 year between operative and conservative treatment but surgery led to faster recovery. All the RCTs aimed to include LDH patients with prolonged pain, exclude absolute operative indications, and use appropriate outcome measures. However, the reporting of RCT characteristics in general was incomplete in all the six RCTs (Table 3). None

of the trials reported patient's path before randomisation or the characteristics of the health care settings. Also, the description of functioning, comorbidity, behavioural and environmental factors, and socioeconomic status were often lacking.

The proportions of operatively treated patients in the RCTs' operative treatment arms varied from 56 to 100% and the percentages of crossing over were high in all RCTs. Bailey et al., Peul et al., and Weinstein et al. reported the results according to the intention-to-treat (ITT) principle and Österman et al. included also an additional on-treatment analysis. Two studies reported cross-over group results separately [2, 4]. Due to high percentages of cross-over, only 16% difference in the shares of patients obtaining operative treatment remained between the index and comparison treatment arms in Weinstein's trial and the corresponding shares were 53, 46, 50, 72 and 61% in the trials by Bailey, Buttermann, Peul, Weber and Österman, respectively. As the primary statistical analysis must be based on groups made comparable by randomization, the poor adherence leads to bias in relation to the original study question [7]. Consequently, none of the studies succeeded in assessing the effectiveness of LDH surgery per se but based on the actual poor adherence to the interventions.

The study question analysis showed that there were no catchment area representative patient groups in any of the RCTs (Fig. 1). Adding the lacking description of patient selection in all the trials, the representativeness of the RCT findings of the catchment areas remains unclear. The study question flow chart also shows that the studies differed from each other in almost all relevant criteria (Fig. 1). Half of the six RCTs included also patients with concomitant lumbar spinal stenosis (LSS). In FinSpine LSS was a concomitant diagnosis in 11.3% of the patients. We find inclusion of LSS patients a strength of the RCTs and registry data. In clinical reality also LSS patients suffer from LDH and understanding that the outcomes in this patient group are similar as the outcomes of patients without LSS gives justification to treat LDH in LSS patients regardless of the underlying degenerative disorders.

The FinSpine registry patients were older than patients in any of the RCTs. Concerning BMI and proportion of female, the registry and the RCTs seemed alike (Table 2). The intensity of leg pain at baseline was higher in the Bailey trial than in the Österman and Peul trials, while other trials did no present leg pain numerically. Baseline leg pain intensity among the FinSpine patients subsided between the trials by Peul and Bailey (Table 5, Fig. 3). The degree of disability estimated with ODI was higher in the Bailey and Weinstein trials than in the Österman trial, while Peul used Roland disability index that was not directly comparable to ODI and numerical information was lacking in two trials [2, 4]. The baseline disability in FinSpine was well comparable with the Bailey and Weinstein trials (Table 5, Fig. 2). At 1 year Bailey trial was the only one to show comparable results to the FinSpine registry concerning leg pain intensity and ODI. All the other trials reporting these measures had lower leg pain intensity and better performance in ODI [5, 6] (Table 4, Figs. 2 and 3).

Participants in randomised trials are selected based on strict inclusion and exclusion criteria, which naturally limits the applicability of the results to ordinary clinical practice. Patients with severe symptoms are more likely to be excluded in the RCTs because of a risk for poor outcome after conservative treatment. We noted that intolerable pain was an exclusion criterion only in the study by Österman et al. which leaves unanswered the question whether patients with intolerable pain were truly randomised in the other studies.

The FinSpine does not allow analysis of patients' paths before surgery, but since the registry covers the whole catchment population, the risk of patient selection is low. The lack of exclusion criteria is found a strength as it results in a heterogenic patient population which is the clinical reality in all spine centres. As conservatively treated patients are not included in the FinSpine, no comparisons between operatively and conservatively treated patients can be made. All extremes concerning the severity of the symptoms are included in the registry. Loss to follow-up may bias the outcome assessment of registry data. However, there is evidence from the Norwegian spine registry that respondents and non-respondents do not differ in outcome, although the non-respondents were often younger and had fewer complications [17].

According to our findings the clinical heterogeneity and incomplete documentation and reporting of the RCTs limits interpretations of effectiveness estimates that would be applicable to clinical reality [7]. The FinSpine is a valuable complement to RCTs, addressing their limitations by providing large-scale, long-term, and real-world data. Integrating findings from both methodologies enhances the robustness of clinical evidence, ensuring interventions are both efficacious in controlled settings and effective in routine clinical practice. Solution to increase applicability even more, would be to implement pragmatic, registry-based randomized controlled trials that integrate the strengths of both RCTs and national registries. This approach addresses the major flaws identified-such as limited generalizability, incomplete reporting, and lack of real-world applicability-by combining rigorous randomization with comprehensive, real-life data collection.

Conclusion

The generalizability and applicability of the results of the RCTs are limited due to several factors. However, the results from RCTs concord with results from the nationwide Finspine and indicate real-life effectiveness of the operative treatments. Our findings show the need for pragmatic RCTs and clinical registry-based data to increase the evidence for decision making in real-life settings. The findings of the present study can be used for planning future RCTs and surgical spine registries on LDH surgery.

Acknowledgements
Not applicable.

Author contributions
All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ida Rantalaiho and Katri Pernaa. The first draft of the manuscript was written by Ida Rantalaiho and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding
This work has been supported with research funding granted by the hospital district of Hlesinki and Uusimaa, Finland, and the funding is appointed to Antti Malmivaara, Grant Number A2500/501.

Data availability
The datasets generated and analysed in the current study can be accessed upon approval of a permit application submitted to the Finnish social and health data permit authority, Findata, in accordance with the Finnish law on the secondary use of health and social data.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 20 May 2025 / Accepted: 7 October 2025
Published online: 04 November 2025

References

1. Bailey CS, Rasoulinejad P, Taylor D, Sequeira K, Miller T, Watson J, et al. Surgery versus conservative care for persistent sciatica lasting 4 to 12 months. N Engl J Med. 2020;382(12):1093–102.
2. Buttermann GR. Treatment of lumbar disc herniation: epidural steroid injection compared with discectomy. A prospective, randomized study. J Bone Joint Surg Am. 2004;86(4):670–9.
3. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JAH, Tans JTJ, et al. Surgery versus prolonged conservative treatment for sciatica. N Engl J Med. 2007;356(22):2245–56.
4. Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. Spine (Phila Pa 1976). 1983;8(2):131–9.
5. Weinstein JN, Tosteson TD, Lurie JD, Tosteson ANA, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA. 2006;296(20):2441–50.
6. Österman H, Seitsalo S, Karppinen J, Malmivaara A. Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. Spine (Phila Pa 1976). 2006;31(21):2409–14.
7. Liu C, Ferreira GE, Abdel Shaheed C, Chen Q, Harris IA, Bailey CS, et al. Surgical versus non-surgical treatment for sciatica: systematic review and meta-analysis of randomised controlled trials. BMJ. 2023;381:e070730.
8. Sharma P, Hughes S, El Haj A, Maffulli N. Expression of the two pore domain potassium channel TREK-1 in human intervertebral disc cells. Curr Stem Cell Res Ther. 2012;7(4):266–71.
9. Longo UG, Papapietro N, Petrillo S, Franceschetti E, Maffulli N, Denaro V. Mesenchymal stem cell for prevention and management of intervertebral disc degeneration. Stem Cells Int. 2012;2012:921053.
10. Longo UG, Petrillo S, Franceschetti E, Maffulli N, Denaro V. Growth factors and anticatabolic substances for prevention and management of intervertebral disc degeneration. Stem Cells Int. 2012;2012:897183.
11. Chen H, Tang T, Xue C, Liu X, Xi Z, Xie L, et al. Exploration and breakthrough in the mode of intervertebral disc cell death may lead to significant advances in treatments for intervertebral disc degeneration. J Orthop Surg Res. 2024;19(1):825.
12. Marjamaa J, Huttunen J, Kankare J, Malmivaara A, Pernaa K, Salmenkivi J, et al. The Finnish spine register (FinSpine): development, design, validation and utility. Eur Spine J. 2023;32:3731–43.
13. Saijonkari M, Aronen P, Booth N. Lannerangan välilevypullistuman hoito ja kuntoutus: Järjestelmällinen katsaus leikkausindikaatioista, hoito-ja kuntoutusmenetelmistä Arviointiryhmä. 2019. Available from: www.summaryx.eu
14. Malmivaara A. Validity and generalizability of findings of randomized controlled trials on arthroscopic partial meniscectomy of the knee. Scand J Med Sci Sports. 2018;28(9):1970–81.
15. Malmivaara A. Assessing validity of observational intervention studies—the benchmarking controlled trials. Ann Med. 2016;48(6):440–3.
16. Malmivaara A. Benchmarking controlled trial–a novel concept covering all observational effectiveness studies. Ann Med. 2015;47(4):332–40.
17. Solberg TK, Sørlie A, Sjaavik K, Nygaard ØP, Ingebrigtsen T. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? Acta Orthop. 2011;82(1):56–63.

Publisher's note