Review Article

AO SPINE

Practice Variation Research in Degenerative Lumbar Disc Surgery: A Literature Review on Design Characteristics and Outcomes

Global Spine Journal
2022, Vol. 12(8) 1841–1851

DOI: 10.1177/21925682211064855
journals.sagepub.com/home/gsj
SAGE

Juliëtte J. C. M. Van Munster, MD[1,2], Vera de Weerdt, MD[3], Ilan J. Y. Halperin, MD[1,2], Amir H. Zamanipoor Najafabadi, MD[2], Peter Paul G. van Benthem, MD, PhD[1], Guus G. Schoonman, MD, PhD[4], Wouter A. Moojen, MD, PhD[2], Wilbert B. van den Hout, PhD[5], Femke Atsma, PhD[6], and Wilco C. Peul, MD, PhD[2]

Abstract

Study Design: Literature review.

Objective: To describe whether practice variation studies on surgery in patients with lumbar degenerative disc disease used adequate study methodology to identify unwarranted variation, and to inform quality improvement in clinical practice. Secondary aim was to describe whether variation changed over time.

Methods: Literature databases were searched up to May 4th, 2021. To define whether study design was appropriate to identify unwarranted variation, we extracted data on level of aggregation, study population, and case-mix correction. To define whether studies were appropriate to achieve quality improvement, data were extracted on outcomes, explanatory variables, description of scientific basis, and given recommendations. Spearman's rho was used to determine the association between the Extreme Quotient (EQ) and year of publication.

Results: We identified 34 articles published between 1990 and 2020. Twenty-six articles (76%) defined the diagnosis. Prior surgery cases were excluded or adjusted for in 5 articles (15%). Twenty-three articles (68%) adjusted for case-mix. Variation in outcomes was analyzed in 7 articles (21%). Fourteen articles (41%) identified explanatory variables. Twenty-six articles (76%) described the evidence on effectiveness. Recommendations for clinical practice were given in 9 articles (26%). Extreme Quotients ranged between 1-fold and 15-fold variation and did not show a significant change over time (rho= −.33, *P*= .09).

Conclusions: Practice variation research on surgery in patients with degenerative disc disease showed important limitations to identify unwarranted variation and to achieve quality improvement by public reporting. Despite the availability of new evidence, we could not observe a significant decrease in variation over time.

Keywords

lumbar, degenerative disc disease, disc herniation, radiculopathy, discectomy, laminectomy

[1]Leiden University Medical Center (LUMC), Leiden, Netherlands
[2]University Neurosurgical Center Holland, Leiden University Medical Center, the Hague Medical Center, and Haga Teaching Hospitals, Leiden and the Hague, the Netherlands
[3]Talma Institution, Vrije Universiteit Amsterdam, the Netherlands & Amsterdam University Medical Centers, Amsterdam, the Netherlands
[4]Elisabeth-TweeSteden Hospital, Tilburg, the Netherlands
[5]Leiden University Medical Center, the Netherlands
[6]Radboud University Medical Center/Radboud Institute for Health Sciences/Scientific Center for Quality of Healthcare (IQ healthcare), Nijmegen, the Netherlands

Corresponding Author:
Juliëtte J. C. M. van Munster, Department of Otorhinolaryngology and Head and Neck Surgery, Leiden University Medical Center, 2300 RC Leiden 2333 ZA, Netherlands.
Email: j.j.c.m.van_munster@lumc.nl

Introduction

Degenerative lumbar spine disorders lead to disability, sick leave, and high societal and healthcare costs.[1,2] The most frequent disorder is degenerative disc disease leading to a herniated disc.[3] In 1934, Mixter and Barr's paper on 19 surgically treated patients with root compression, officially opened the era of spine surgery.[4] In the last decades, the number of lumbar spine procedures for lumbar back pain and leg pain increased substantially and large variation in surgical rates was observed between and within regions.[5-8]

Unwarranted variation in surgical rates is variation that cannot be explained by differences in patient needs and preferences.[9] Hence, it can be driven by the lack of high-quality evidence on indications for surgery or differences in surgeons' beliefs about the effectiveness of procedures. High-quality research on the effectiveness of surgical treatment in degenerative lumbar disc disease is still lacking for some procedures, but has increased significantly in the last decades.[10] Implementation of Evidence-Based Medicine (EBM) to improve healthcare can be achieved by the development of clinical guidelines.[11] Subsequently, appropriate studies on unwarranted variation can be used as feedback to clinicians and policy makers to improve implementation of these guidelines. Public reporting of these studies can be a first step toward change and might close the loop between EBM and clinical practice.[12,13] However, analyzing and explaining practice variation is challenging because multiple factors influence variation in surgical rates.[9] Previously, it was described that clinical audits or practice variation research must: 1. Select a diagnosis, 2. Provide a scientific basis to demonstrate the gap between actual and desired practice, and 3. Define warranted and unwarranted use of the target outcome.[13-16] Ideally, the article should investigate the causes of practice variation to identify specific areas for improvement and provide recommendations for clinical practice.[13,17]

We aimed to describe whether practice variation studies on lumbar disc surgery in patients with lumbar degenerative disc disease used adequate study methodology to identify unwarranted variation and to inform quality improvement in clinical practice. Secondary, we were interested whether variation changed over time. We hypothesize that the availability of new evidence and guidelines combined with the attention for unwarranted variation in spine surgery of the past decades led to lower variation in more recent years.[18]

Methods

This literature review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[19] We did not register a study protocol.

Databases and Selection Process of Studies

A literature search was performed by a trained librarian on PubMed, Embase, and Web of Science on May 04th, 2021 to identify articles on practice variation in lumbar disc surgery for degenerative disc disease in adults. Title and abstract screening as well as full text screening was performed by two reviewers independently (JM, VW). A third and fourth reviewer (FA, WP) were consulted in case of conflicts between the two reviewers. Records were managed through Rayyan;[20] specific software for managing bibliographies.

Search Strategy and Inclusion Criteria

We focused on peer-reviewed studies on practice variation in lumbar disc surgery for degenerative disease. Our search strategy consisted of three main concepts and variations thereof: "spinal diseases or low back pain," "spine surgery," and "practice variation or small area analysis." The full search strategy can be found in the supplemental files. Articles that focused on lumbar disc surgery without specification of the disease in the methods were also included. Hereby, we aimed to identify on which diagnoses these articles focused and which diagnosis codes and procedure codes were included. We excluded articles on lumbar disc surgery for malignancies, traumatic fractures, spinal deformities, and congenital diseases, as the pathophysiological mechanisms differ from degenerative lumbar disc disease. Furthermore, we excluded articles that analyzed practice variation in cervical or thoracic spine surgery only. Also, survey studies using case scenarios were excluded. Lastly, articles were excluded if no full text was available and if articles were not written in English.

Data-Extraction

Two reviewers (JM and VW) extracted data on the study characteristics. Indicators for appropriate study design to identify unwarranted variation and optimal study design to achieve quality improvement were based on previous research and frameworks.[13,14,17,21-23] First, characteristics were extracted to describe whether study design was appropriate to identify unwarranted variation: the level of aggregation, study population (inclusion and exclusion criteria), description and selection of the diagnosis or diagnosis group, and variables used for case-mix correction. Second, characteristics were extracted to describe whether study design was optimal to achieve quality improvement: variation in clinical outcomes, analyzed explanatory variables (other than the variables adjusted for), scientific basis for treatment effectiveness described (i.e., practices compared against clinical guidelines), and if recommendations for clinical practice or future research were given. Third, we described coding used for the procedures and the diagnoses. Lastly, we described whether significant variation was concluded by the authors (yes vs. no) and the Extreme Quotient (EQ, highest/lowest surgical rate).

Analysis

Spearman's rho was used to determine the association between time (year of publication) and study outcome (EQ). We hypothesized lower variation in more recent years due to the

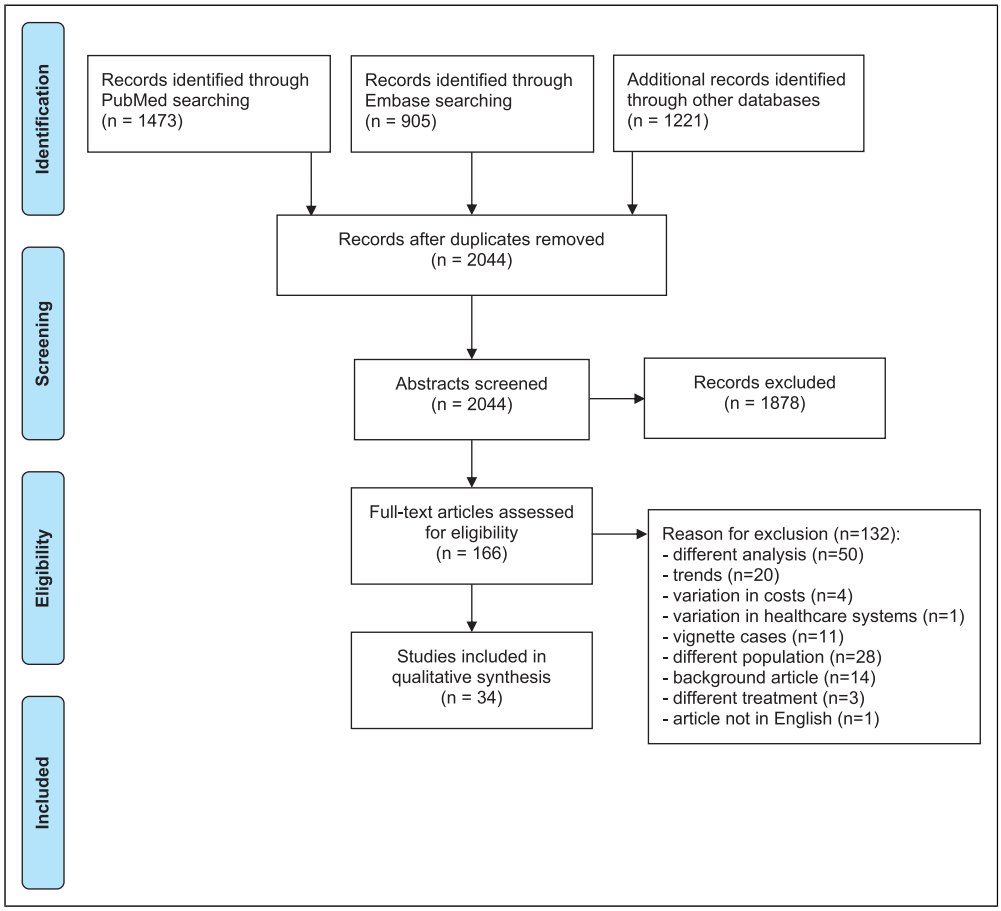

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram on study selection.

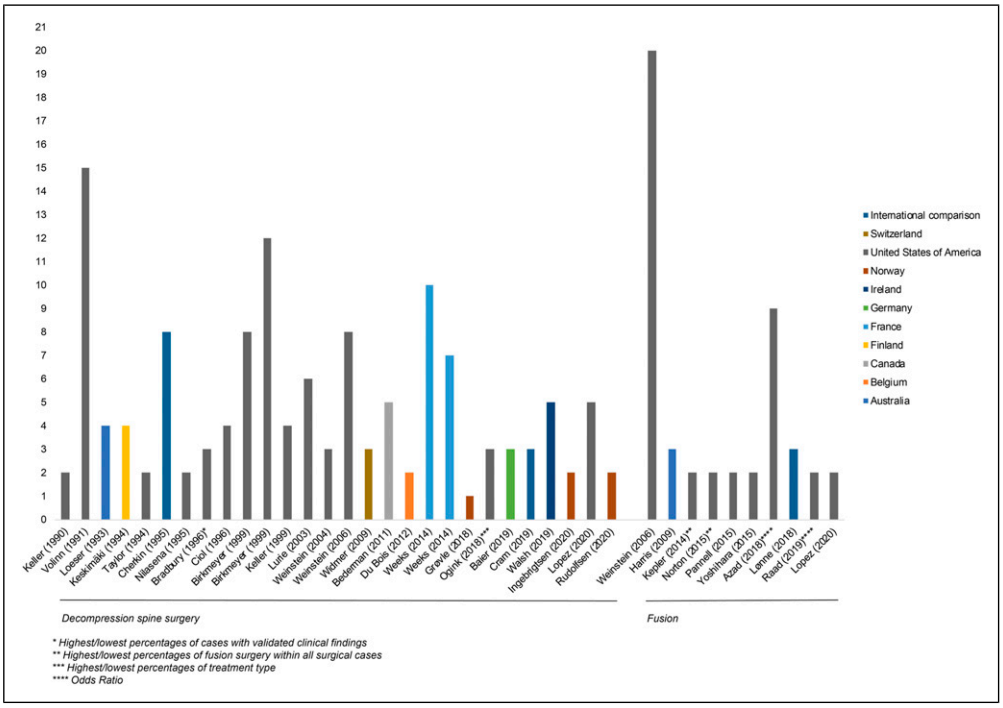

Figure 2. Numbers of publications and Extreme Quotient over time by type of surgery.

Table 1. Article Characteristics Important to Identify Unwarranted Variation.

Author (Year)	EQ[a]	Diagnosis specified	Diagnosis selected	Exclusion prior surgery	Case-mix correction				Level of aggregation
					Referrals	Severity	Preferences	Other	
Ogink (2018)	NA[b]	Specific	Yes	No	Yes	Yes	No	No	Individual
Keller (1990)	3	No	No	No	Yes	No	No	Yes	Hospital or hospital service area
Loeser (1993)	4	Multiple	Yes	No	Yes	No	No	No	
Keskimaki (1994)	4	No	No	No	Yes	No	No	Yes	
Bradbury (1996)	NA	Multiple	Yes	No	No	Yes	No	No	
Birkmeyer (1999)	8	Multiple	Yes	No	Yes	No	No	Yes	
Keller (1999)	4	Multiple	No	No	Yes	No	No	Yes	
Lurie (2003)	6	No	No	No	Yes	No	No	Yes	
Weinstein (2004)	5	Multiple	No	No	Yes	No	No	Yes	
Weinstein (2006)	8	No	Yes	No	Yes	No	No	Yes	
Widmer (2009)	3	Specific	No	No	Yes	No	No	Yes	
Baier (2019)	3	No	Yes	No	No	No	No	Yes	
Walsh (2019)	5	Multiple	Yes	No	Yes	No	No	Yes	
Ingebrigtsen (2020)	2	Multiple	Yes	Yes	Yes	No	No	Yes	
Volinn (1991)	15	Multiple	Yes	No	Yes	No	No	Yes	Regional
Taylor (1994)	2	Multiple	Yes	No	No	No	No	Yes	
Nilasena (1995)	2	Multiple	Yes	No	Yes	No	No	Yes	
Ciol (1996)	4	Specific	No	No	Yes	No	No	Yes	
Harris (2009)	3	Multiple	Yes	No	No	No	No	No	
Bederman (2011)	5	Multiple	Yes	No	Yes	No	No	Yes	
Du Bois (2012)	2	Multiple	No	No	Yes	No	No	Yes	
Kepler (2014)	NA	Specific	Yes	No	No	No	No	No	
Weeks (2014)	7	Multiple	Yes	No	No	No	No	Yes	
Yoshihara (2015)	2	Specific	Yes	No	No	No	No	No	
Norton (2015)	NA	Specific	Yes	No	No	No	No	No	
Pannell (2015)	2	No	No	No	Yes	No	No	Yes	
Grøvle (2018)	1	Multiple	Yes	No	Yes	No	No	No	
Azad (2018)	NA	Specific	Yes	Yes	Yes	No	No	Yes	
Raad (2019)	NA	Specific	Yes	No	Yes	No	No	Yes	
Lopez (2020)	5	No	No	No	No	No	No	No	
Rudolfsen (2020)	2	Multiple	Yes	Adjusted	Yes	Yes	No	Yes	
Cherkin (1995)	8	Multiple	Yes	No	Yes	No	No	No	National
Lønne (2018)	3	Specific	Yes	Yes	No	No	No	No	
Cram (2019)	3	No	No	Yes	Yes	No	No	Yes	

[a]Extreme Quotient (Highest/lowest surgical rate).
[b]Not applicable.

increased number of studies on effectiveness of surgical procedures. For this analysis, studies that did not describe the highest and lowest population-based surgical rates or EQ and studies that focused on fusion surgery only were excluded.

Results

The process of study identification and selection is presented in Figure 1. In total, 34 articles published between 1990 and 2020 were included. Nineteen articles (56%) investigated practice variation in the USA. Most articles (n = 30) used administrative healthcare databases for the analyses. Three articles used national spine registries for data collection,[24-26] and one article used hospital records.[27] In all but three articles,[25,28,29] significant variation of surgical rates was concluded by the authors. Highest and lowest population-based surgical rates were described in 26 articles (76%). The EQ ranged from 1-fold to 15-fold in surgical rates. We observed a median EQ of 4-fold variation (Interquartile Range 2.0–7.3). We did not observe a significant decrease in EQ over time (Figure 2, rho = −.24, P = .2). If articles were depicted twice, practice variation in different type of procedures was analyzed.

Identifying Unwarranted Variation

Characteristics important to identify unwarranted variation were described in Table 1. One article investigated differences

Table 2. Investigated Population And Included Diagnosis Codes.

Author (Year)	Population	Diagnosis	721	722	724	738	756	846	847	Other	NA[a]
Keller (1990)	Hospitalized	Not specified						X			
Keskimaki (1994)	All	Not specified									X
Lurie (2003)	Medicare, >65y	Not specified									X
Pannell (2015)	Private insured	Not specified									X
Baier (2019)	All	Not specified								X[b]	
Cram (2019)	Age>18y[c]	Not specified									X
Lopez (2020)	Medicare, >65y	Not specified									X
Volinn (1991)	All	LBP[d]	X	X	X	X	X		X		
Loeser (1993)	All	LBP, sciatica	X	X	X	X	X		X		
Taylor (1994)	Hospitalized, >20y	LBP, sciatica	X	X	X	X	X	X	X		
Cherkin (1995)	All	Back pain	X	X	X	X	X	X	X	X[e]	
Nilasena (1995)	Age>65	Back problems	X	X	X	X	X	X	X	X[f,g]	
Weinstein (2006)	Medicare, >65y	Back problems	X	X	X	X	X	X	X	X[g]	
Du Bois (2012)	All	LBP									X
Walsh (2019)	Hospitalized, >18y	Degenerative LBP								X[b]	
Bradbury (1996)	All	Spine problems									X
Birkmeyer (1999)	Medicare, >65y	LDH[h], LSS	X	X	X	X	X	X	X	X[f,g]	
Keller (1999)	Hospitalized	Sciatica									X
Weinstein (2004)	Medicare, >65y	DDLS[i]									X
Harris (2009)	All	DDLS									X
Bederman (2011)	Age>50	DDLS	X		X	X					
Weeks (2014)	Age>35	Spine problems								X[b]	
Yoshihara (2015)	Hospitalized, >18y	DDLS								X[j]	
Ingebrigtsen (2020)	Public hospital	DDLS, fractures								X[b]	
Rudolfsen (2020)	Age>16y	DDLS									X
Ciol (1996)	Medicare, >65y	LSS[k]									X
Widmer (2009)	Hospitalized, >18y	LSS									X
Grøvle (2018)	Public hospital, >18y	LSS, spondylosis								X[b]	
Lønne (2018)	Normal BMI, <50y	LSS								X[l]	
Ogink (2018)	Age>18y	LSS									X
Raad (2019)	Insured, >40<65y	LSS		X							
Kepler (2014)	All	Spondylolisthesis				X					
Norton (2015)	All	Spondylolisthesis				X					
Azad (2018)	Insured, age<65y	Spondylolisthesis				X	X				

[a]Not applicable: no description of included codes.
[b]ICD-10 coding.
[c]years.
[d]Low back pain.
[e]National coding Norway and Canada.
[f]739.
[g]996.
[h]Lumbar Disc Herniation.
[i]Degenerative Disease of the Lumbar Spine.
[j]772.
[k]Lumbar Spinal Stenosis.
[l]Data from national spine registries.

between surgeons in the percentage of surgically treated patients.[30] Five studies investigated differences in percentages of fusion in patients that underwent decompressive spine surgery.[24,31-34] Lastly, one article investigated differences between hospitals in validating clinical findings for surgery.[27] The evaluated articles described analyses of surgical rates on different levels of aggregation, but not one article described differences in first line treatment. Inclusion and exclusion criteria varied between articles. Twenty-six articles (76%) selected a diagnosis or diagnosis group for their analyses. Four articles (12%) mentioned that prior surgery cases were excluded,[24,25,32,35] and one article (3%) adjusted for prior

surgery.[26] Ten articles (29%) defined criteria for age, but cut-offs varied between articles (Table 2). Moreover, the age criteria did not always match with the investigated diagnosis.

Case-mix correction was performed in 23 articles (68%)[6,25,29,30,32,35-51] (Figure 3). No article investigated timing of surgery or adjusted for disease severity to define unwarranted variation in treatment choice. In thirteen articles (38%), adjustment for referral cases was accomplished by analyzing practice variation on the level of Hospital Service Areas (HSAs). However, HSAs were defined in different ways. For example, Keller et al. defined spine service area, using discharges for spine problems only,[36] whereas other articles based the HSAs on neurosurgery and major cardiovascular procedures, or on all discharges.

Coding of Procedures and Diagnoses

Used coding is described in Tables 2 and 3. Specific codes varied more widely (Supplemental Tables 1 and 2). Twelve studies (35%) based their analysis on procedure codes without matching diagnosis codes,[26,29,35,36,41,44,45,47,49,51-54] and fifteen studies (44%) based their analysis on matching both procedure codes and diagnosis codes.[25,28,37,40-42,46,48,50,55,56] One study based the analysis on diagnosis codes.[38]

The differences in definition of the diagnosis could partly be explained by focusing on different diagnosis groups. However, if similar coding for the diagnoses was used, variation in definition of the disease occurred. For example, Nilasena et al.[40] described a long list of problems including "nonspecific backache" and "instability" for the same codes that Birkmeyer et al.[42] defined as "spinal stenosis or lumbar disc herniation." Similarly, if approximately the same definition of the disease (i.e., lumbar degenerative disease) was used, coding varied widely. Ten articles (29%) included diagnoses for low back pain without sciatica.[29,36,38,40,42,46,48,51,52,55] Investigated procedures and procedure codes varied as well between articles (Table 3). Twenty-three articles (68%) investigated practice variation in discectomies, twenty articles (59%) investigated laminectomies, and twenty-seven articles (79%) investigated fusion. Furthermore, four articles included the code "Lysis of adhesions of spinal cord"[38,42,46,50]; three articles included the code "Internal fixation of bone"[38,42,46]; two articles included the code for "Insertion of spinal disc prosthesis"[50,56]; and two articles included codes for refusion.[50,52]

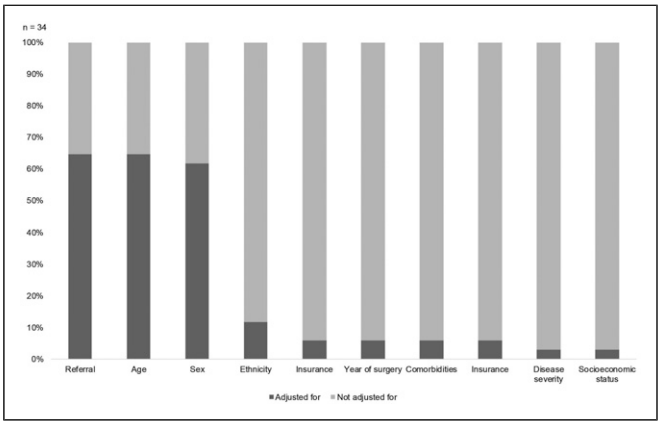

Figure 3. Adjustment for case-mix in 34 articles analyzing practice variation in degenerative lumber spine surgery.

Study Design Optimal for Quality Improvement

Study characteristics important for quality improvement were described in Table 4. Regional variation in patient's outcomes was analyzed in 7 articles (21%),[24,26,28,31,35,43,48] of which 5 articles were published in the most recent years (2018-2020). Fourteen articles (41%) tried to identify variables that explain the variation in surgical rates.[6,26,27,29,30,32,37,40,44,46,48,51,52,55] Patient factors, surgeon factors, and supply factors were included in these analyses, but also the use of spinal imaging,[44] and competition between hospitals.[29] Most articles (n = 25, 78%) described the level of evidence for effectiveness, but in 17 articles it was described that evidence of effectiveness was lacking. Recommendations for future research were given in 25 articles (74%) and recommendations for clinical practice were given in 9 articles (26%).

Discussion

Although we observed large differences in variation between studies, we did not observe a decrease in practice variation in surgical rates for lumbar degenerative disc disease over time between 1990 and 2020. The most recent studies on variation implied that the practice variation is still problematic, despite the fact that evidence on effectiveness and timing of surgery improved in the last decades.[57-60] However, the largest variations in surgical rates were described in one of the very first articles for both general spine surgery and fusion.[46,61] Furthermore, we observed important limitations in the study design of practice variation research on lumbar disc surgery in patients with degenerative disc disease. Not all articles used adequate study design to identify unwarranted variation and to be able to inform improve quality improvement. Moreover, we observed substantial heterogeneity in study methodology, hampering the comparison of practice variation studies over time and between regions and countries.

Not all study designs were appropriate to identify unwarranted variation. First, the diagnosis group of interest was not clearly defined in all studies. Moreover, most articles that did define the diagnosis group included multiple diagnoses. We were surprised by these findings since the indication for surgical treatment depends on the patient's diagnosis. Second, none of the studies included timing of surgical treatment in the analysis, while this is an important indication for surgery in patients suffering sciatica due to lumbar disc herniation.[57] Lastly, not all studies adjusted for relevant case-mix factors. Future practice variation studies should specify both the

Table 3. Included Procedure Codes.

Author (Year)	Procedure	03.0	03.1	78	80.5	81	84.6	Other	NA[a]
Keller (1990)	Disc excision				X				
Bradbury (1996)	Disc excision				X				
Volinn (1991)	Low back surgery	X			X	X			
Nilasena (1995)	Low back surgery	X			X	X			
Cherkin (1995)	Back surgery	X			X	X		X[b,c]	
Loeser (1993)	Lumbar spine surgery	X			X	X			
Keskimaki (1994)	Lumbar spine surgery								X
Taylor (1994)	Lumbar spine surgery	X	X	X	X	X	X		
Ciol (1996)	Lumbar spine surgery	X	X	X	X	X			
Birkmeyer (1999)	Lumbar spine surgery	X	X	X	X	X			
Keller (1999)	Lumbar spine surgery								X
Bederman (2011)	Lumbar spine surgery							X[d]	
Du Bois (2012)	Lumbar spine surgery							X[e]	
Norton (2015)	Lumbar spine surgery					X			
Ingebrigtsen (2020)	Lumbar spine surgery							X[f]	
Lurie (2003)	Spine surgery							X[g]	
Weinstein (2004)	Spine surgery	X	X		X	X			
Weinstein (2006)	Spine surgery	X	X	X	X	X			
Widmer (2009)	Spine surgery							X[h]	
Weeks (2014)	Spine surgery	X	X		X	X	X		
Baier (2019)	Spine surgery							X[i]	
Cram (2019)	Spine surgery							X[j]	
Rudolfsen (2020)	Spine surgery								X
Grøvle (2018)	Spinal stenosis surgery							X[f]	
Ogink (2018)	Spinal stenosis surgery								X
Harris (2009)	Fusion							X[k]	
Kepler (2014)	Fusion							X[g]	
Pannell (2015)	Fusion								X
Yoshihara (2015)	Fusion, disc replacement					X	X		
Azad (2018)	Fusion					X		X[g]	
Lønne (2018)	Fusion							X[l]	
Raad (2019)	Fusion	X			X	X			
Walsh (2019)	Fusion, decompression	X			X	X		X[j]	
Lopez (2020)	Fusion, discectomy							X[g]	

[a]Not applicable (no description).
[b]National coding Canada and Norway.
[c]Loeser–Volinn algorithm used in Australia.
[d]Canadian Classification of Procedures.
[e]Belgian nomenclature.
[f]NOMESCO classification of surgical procedures.
[g]Current Procedure Technology.
[h]CHOP treatment classification.
[i]Operation and procedure codes.
[j]ICD-10.
[k]Medicare Australia Codes.
[l]Data from national spine registries.

diagnosis and procedure, exclude repeat surgery, and adjust for relevant case-mix, including timing of surgery and severity of the disease if possible.

The limitations in methodology might partly be caused by the limitations of administrative healthcare databases. These databases have important advantages and disadvantages. These databases enable investigation of large geographic areas and coverage over multiply years, which is an advantage for measuring practice variation.[62] Although quality of administrative databases improved over the last decades, they also have some drawbacks for measuring practice variation in spine surgery. For timing of surgery, linking between primary care and hospital databases is necessary, which is not always possible. Linking these databases will enable analysis of the

Table 4. Article Characteristics Important to Achieve Quality Improvement.

Author (Year)	Variation in outcomes analyzed	Explanatory variables analyzed[a]	Evidence for effectiveness described	Recommendations clinical practice	Recommendations future research	Level of aggregation
Ogink (2018)	No	Yes	Yes	Yes	No	Individual
Keller (1990)	No	No	NA[b]	No	Yes	Hospital or hospital service area
Loeser (1993)	No	Yes	No	No	No	
Keskimaki (1994)	No	No	No	No	No	
Bradbury (1996)	No	Yes	No	No	Yes	
Birkmeyer (1999)	No	No	NA	No	Yes	
Keller (1999)	Yes	Yes	NA	Yes	Yes	
Lurie (2003)	No	Yes	No	No	Yes	
Weinstein (2004)	No	No	NA	Yes	Yes	
Weinstein (2006)	No	Yes	NA	No	Yes	
Widmer (2009)	No	No	NA	Yes	Yes	
Baier (2019)	No	Yes	Yes	Yes	Yes	
Walsh (2019)	No	Yes	NA	No	Yes	
Ingebrigtsen (2020)	No	No	Yes	Yes	No	
Volinn (1991)	No	No	NA	No		Regional
Taylor (1994)	No	Yes	NA	No	Yes	
Nilasena (1995)	No	No	NA	No	Yes	
Ciol (1996)	No	No	NA	No	No	
Harris (2009)	No	Yes	Yes	No	Yes	
Bederman (2011)	Yes	No	Yes	Yes	Yes	
Du Bois (2012)	No	No	NA	No	Yes	
Kepler (2014)	No	No	No	No	Yes	
Weeks (2014)	No	No	NA	No	No	
Yoshihara (2015)	No	No	Yes	No	Yes	
Norton (2015)	No	No	NA	No	Yes	
Pannell (2015)	No	No	No	No	No	
Grøvle (2018)	Yes	Yes	NA	No	Yes	
Azad (2018)	No	Yes	No	No	No	
Raad (2019)	Yes	No	Yes	No	Yes	
Lopez (2020)	No	No	No	No	Yes	
Rudolfsen (2020)	Yes	Yes	Yes	Yes	Yes	
Cherkin (1995)	No	Yes	NA	No	Yes	National
Lønne (2018)	Yes	No	NA	No	Yes	
Cram (2019)	Yes	No	NA	Yes	No	

[a]Other than the variables adjusted for.
[b]Not Applicable: article described lack of evidence on effectiveness.

full care path, including nonoperative guidance by the general practitioner and physiotherapist, and use timing of surgery as a quality indicator. Additionally, not all the drivers of practice variation can be measured as case-mix variables due to the lack of availability in databases.

The differences in coding within similar diagnosis and procedure groups hamper comparison between articles. This might be caused by the focus of the research question. For example, some articles specifically focused on variation in fusion procedures for degenerative spondylolisthesis. It is no surprise that different diagnosis codes were included in these studies compared to studies that focus on degenerative disc disease. However, methodology differed within studies investigating similar diagnosis groups as well. Another reason for this finding might be the differences in available codes between administrative databases. Moreover, coding will depend on provider registration and interpretation of medical coders. Standardized terminology and coding based on ICD codes within all countries can also improve the quality of comparisons within and between international databases.

Lastly, most administrative healthcare databases do not include clinical outcomes. Reporting on variation in clinical outcomes as a result of variation in clinical practice can contribute to the intrinsic motivation of physicians to deliver

the best care for their patients and thereby facilitate the quality improvement process.[17] Not all studies used optimal study design to identify areas for quality improvement and close the loop between EBM and clinical practice. Although public reporting might be the first step towards change,[12] optimal study design can improve the impact of an article.[17] Future research should ideally compare practices against clinical practice guidelines, include outcome variables, and give recommendations for clinical practice. Advancements in the quality and comprehensiveness of administrative databases and linking between clinical outcome databases and administrative databases will facilitate the possibility to use methodology important for quality improvement.

To the best of our knowledge, this is the first review that described differences and limitations in methods for practice variation research in lumbar degenerative disc disease. A strength of this review was the systematic search and data selection providing a comprehensive overview of all relevant methodological and clinical aspects regarding practice variation in lumbar degenerative spine surgery. Furthermore, not only spinal neurosurgeons, but also a neurologist and methodological experts were involved in our team giving knowledge on clinical and methodological features.

Our study has limitations as well. First, we focused on variation in surgical treatment, whereas variation in conservative treatments, such as physiotherapy or the use of opioids and variation in outcomes are important areas for quality improvement as well. Second, the number of papers was too small to make a proper comparison between the EQ and methodological features. For example, case-mix correction adjusts for the effect of patient characteristics on treatment choice, potentially leading to lower variation in surgical rates. Third, we only included peer-reviewed articles, missing articles published by national institutes. This includes publications from the Dartmouth Institute of Healthcare,[63] although data were described in peer-reviewed articles as well.[42,44-46] Lastly, we were unable to describe all country specific regulations of spine care, such as specialized spine clinics and the presence or absence of health care insurance, which might contribute to regional differences and must be included in the analysis in order to identify unwarranted variation. Therefore, the non-significant change in practice variation over time is only an indication and must be interpreted with caution.

Conclusions

Practice variation research on lumbar disc surgery in patients with degenerative disc disease showed important limitation in used methodologies that contribute to the possibility of identifying unwarranted variation and improve quality in clinical practice. Furthermore, significant heterogeneity in study designs was observed. This finding could not fully be explained by differences in investigated diagnosis groups. Despite the availability of new evidence, we did not find evidence of a clear decrease in variation over time. However, questions might be raised about the comparability of these studies. Future practice variation studies should specify both the diagnosis and procedure, exclude repeat surgery, and adjust for relevant case-mix, including timing of surgery and severity of the disease if possible. Furthermore, specific national regulations of spine care should be included in the analysis. Lastly, future research should ideally compare practices against clinical practice guidelines, include outcome variables and give recommendations for clinical practice. Hopefully, future practice variation studies will identify areas for quality improvement and close the loop between EBM and clinical practice to improve patient outcomes.

Acknowledgements

We thank Jan Schoones from the Walaeus Library (Leiden University Medical Center) for designing and performing the literature search

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Juliëtte van Munster https://orcid.org/0000-0002-1670-1800

Supplemental Material

Supplemental material for this article is available online.

References

1. 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*. 2011;36:1050-1058. doi:10.1097/BRS.0b013e3181e70488.
2. Torio CM, Andrews RM. National inpatient hospital costs: the most expensive conditions by payer, 2011: statistical brief #160. In: *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2006.
3. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ*. 2007;334:1313-1317. doi:10.1136/bmj.39223.428495.
4. Mixter WJ, Barr JS. Rupture of the intervertebral disc with involvement of the spinal canal. *N Engl J Med*. 1934;211:210-215. doi:10.1056/nejm193408022110506.
5. Jordan J, Konstantinou K, O'Dowd J. Herniated lumbar disc. *Clin Evid*. 2011;2011:1118.
6. Pannell WC, Savin DD, Scott TP, Wang JC, Daubs MD. Trends in the surgical treatment of lumbar spine disease in the United

States. *Spine J*. 2015;15:1719-1727. doi: 10.1016/j.spinee.2013.10.014.
7. Davis H. Increasing rates of cervical and lumbar spine surgery in the United States, 1979-1990. *Spine*. 1994;19:1117-1122; discussion 1123-1114. doi:10.1097/00007632-199405001-00003.
8. Atkinson L, Zacest A. Surgical management of low back pain. *Med J Aust*. 2016;204:299-300. doi:10.5694/mja16.00038.
9. Birkmeyer JD, Reames BN, McCulloch P, Carr AJ, Campbell WB, Wennberg JE. Understanding of regional variation in the use of surgery. *Lancet*. 2013;382:1121-1129.
10. Jacobs WCH, Rubinstein SM, Willems PC, et al. The evidence on surgical interventions for low back disorders, an overview of systematic reviews. *Eur Spine J*. 2013;22:1936-1949. doi:10.1007/s00586-013-2823-4.
11. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet*. 2003;362:1225-1230. doi: 10.1016/S0140-6736(03)14546-1.
12. Westert GP, Groenewoud S, Wennberg JE, et al. Medical practice variation: public reporting a first necessary step to spark change. *Int J Qual Health Care*. 2018;30:731-735. doi:10.1093/intqhc/mzy092.
13. Dazley JM, Cha TD, Harris MB, Bono CM. Closing the loop between evidence-based medicine and care delivery: a possible role for clinical audits in spinal surgery. *Spine J*. 2013;13:1951-1957. doi:10.1016/j.spinee.2013.03.055.
14. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. *J Eval Clin Pract*. 2011;17:671-677. doi:10.1111/j.1365-2753.2011.01689.x.
15. Ivers NM, Sales A, Colquhoun H, et al. No more 'business as usual' with audit and feedback interventions: towards an agenda for a reinvigorated intervention. *Implement Sci*. 2014;9:14. doi: 10.1186/1748-5908-9-14.
16. Strauss ES, Tetroe J, Graham ID. *Knowledge Translation in Health Care: Moving from Evidence to Practice*, 2nd ed. John Wiley & Sons, Ltd; 2013.
17. Karnon J, Partington A, Horsfall M, Chew D. Variation in clinical practice: a priority setting approach to the staged funding of quality improvement. *Appl Health Econ Health Pol*. 2016;14:21-27. doi:10.1007/s40258-015-0160-y.
18. Goodwin JS. Tracking medicine: a researcher's quest to understand health care by John E. Wennberg. *Am J Epidemiol*. 2011;174:252. doi:10.1093/aje/kwr152.
19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097. doi:10.1371/journal.pmed.1000097.
20. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev*. 2016;5:210. doi:10.1186/s13643-016-0384-4.
21. Sepucha K, Ozanne E, Mulley AG Jr. Doing the right thing: systems support for decision quality in cancer care. *Ann Behav Med*. 2006;32:172-178. doi:10.1207/s15324796abm3203_2.
22. Vink MDH, de Bekker PJGM, Koolman X, et al. Design characteristics of studies on medical practice variation of caesarean section rates: a scoping review. *BMC Pregnancy Childbirth*. 2020;20:478. doi:10.1186/s12884-020-03169-3.
23. Partington A, Chew DP, Ben-Tovim D, Horsfall M, Hakendorf P, Karnon J. Screening for important unwarranted variation in clinical practice: a triple-test of processes of care, costs and patient outcomes. *Aust Health Rev*. 2017;41:104-110. doi:10.1071/ah15101.
24. Lønne G, Fritzell P, Hägg O, et al. Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries. *Spine J*. 2019;19:41-49. doi:10.1016/j.spinee.2018.05.028.
25. Ingebrigtsen T, Balteskard L, Guldhaugen KA, et al. Behandlingsrater for ryggkirurgi i Norge og Helse Nord 2014-18. *Tidsskrift for Den norske legeforening*. 2020;140:1-11. doi:10.4045/tidsskr.20.0313.
26. Rudolfsen JH, Solberg TK, Ingebrigtsen T, Olsen JA. Associations between utilization rates and patients' health: a study of spine surgery and patient-reported outcomes (EQ-5D and ODI). *BMC Health Serv Res*. 2020;20:135. doi:10.1186/s12913-020-4968-2.
27. Bradbury RC, Golec JH, Steen PM. Occurrence of validating clinical findings in the hospital record for appendectomy, cholecystectomy and intervertebral disc excision patients: comparisons among hospitals and among surgeons. *Health Serv Manag Res*. 1996;9:34-44. doi:10.1177/095148489600900104.
28. Grøvle L, Fjeld OR, Haugen AJ, et al. The rates of LSS Surgery in Norwegian public hospitals: a threefold increase from 1999 to 2013. *Spine*, 44; 2019:e372-e378. doi:10.1097/brs.0000000000002858.
29. Baier N, Sax L-M, Sundmacher L. Trends and regional variation in rates of orthopaedic surgery in Germany: the impact of competition. *Eur J Health Econ*. 2019;20:163-174. doi:10.1007/s10198-018-0990-2.
30. Ogink PT, van Wulfften Palthe O, Teunis T, et al. Practice variation among surgeons treating lumbar spinal stenosis in a single institution. *Spine*. 2019;44:510-516. doi:10.1097/brs.0000000000002859.
31. Raad M, Reidler JS, El Dafrawy MH, et al. US regional variations in rates, outcomes, and costs of spinal arthrodesis for lumbar spinal stenosis in working adults aged 40-65 years. *J Neurosurg Spine*. 2019;30:83-90. doi:10.3171/2018.5.Spine18184.
32. Azad TD, Vail D, O'Connell C, Han SS, Veeravagu A, Ratliff JK. Geographic variation in the surgical management of lumbar spondylolisthesis: characterizing practice patterns and outcomes. *Spine J*. 2018;18:2232-2238. doi:10.1016/j.spinee.2018.05.008.
33. Norton RP, Bianco K, Klifto C, Errico TJ, Bendo JA. Degenerative spondylolisthesis. *Spine*. 2015;40:1219-1227. doi:10.1097/brs.0000000000000987.
34. Kepler CK, Vaccaro AR, Hilibrand AS, et al. National trends in the use of fusion techniques to treat degenerative spondylolisthesis. *Spine*. 2014;39:1584-1589. doi:10.1097/brs.0000000000000486.
35. Cram P, Landon BE, Matelski J, et al. Utilization and outcomes for spine surgery in the United States and Canada. *Spine*. 2019;44:1371-1380. doi:10.1097/brs.0000000000003083.
36. Keller RB, Soule DN, Wennberg JE, Hanley DF. Dealing with geographic variations in the use of hospitals. The experience of

the maine medical assessment foundation orthopaedic study group. *J Bone Joint Surg*. 1990;72:1286-1293.
37. Volinn E, Turczyn KM, Loeser JD. Theories of back pain and health care utilization. *Neurosurg Clin*. 1991;2:739-748.
38. Taylor VM, Deyo RA, Cherkin DC, Kreuter W. Low back pain hospitalization. *Spine*. 1994;19:1207-1212; discussion 1213. doi:10.1097/00007632-199405310-00002.
39. Keskimäki I, Aro S, Teperi J. Regional variation in surgical procedure rates in Finland. *Scand J Soc Med*. 1994;22:132-138. doi:10.1177/140349489402200209.
40. Nilasena DS, Vaughn RJ, Mori M, Lyon JL. surgical trends in the treatment of diseases of the lumbar spine in utah's medicare population, 1984 to 1990. *Med Care*. 1995;33:585-597. doi:10.1097/00005650-199506000-00002.
41. Ciol MA, Deyo RA, Howell E, Kreif S. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc*. 1996;44:285-290. doi:10.1111/j.1532-5415.1996.tb00915.x.
42. Birkmeyer NJ, Weinstein JN. Medical versus surgical treatment for low back pain: evidence and clinical practice. *Effect Clin Pract*. 1999;2:218-227.
43. Keller RB, Atlas SJ, Soule DN, Singer DE, Deyo RA. Relationship between rates and outcomes of operative treatment for lumbar disc herniation and spinal stenosis. *J Bone Joint Surg*. 1999;81:752-762. doi:10.2106/00004623-199906000-00002.
44. Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine*. 2003:28;616-620. 10.1097/01.Brs.0000049927.37696.
45. Weinstein JN, Bronner KK, Morgan TS, Wennberg JE. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. *Health Aff*. 2004;23:VAR81-VAR89. Suppl Variation. doi:10.1377/hlthaff.var.81.
46. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES. United States' trends and regional variations in lumbar spine surgery: 1992-2003. *Spine*. 2006;31:2707-2714. doi:10.1097/01.brs.0000248132.15231.fe.
47. Widmer M, Matter P, Staub L, Schoeni-Affolter F, Busato A. Regional variation in orthopedic surgery in Switzerland. *Health Place*. 2009;15:791-798. doi:10.1016/j.healthplace.2008.12.009.
48. Bederman SS, Coyte PC, Kreder HJ, Mahomed NN, McIsaac WJ, Wright JG. Who's in the driver's seat? The influence of patient and physician enthusiasm on regional variation in degenerative lumbar spinal surgery. *Spine*. 2011;36:481-489. doi:10.1097/brs.0b013e3181d25e6f.
49. Du Bois M, Szpalski M, Donceel P. A decade's experience in lumbar spine surgery in Belgium: sickness fund beneficiaries, 2000-2009. *Eur Spine J*. 2012;21:2693-2703. doi:10.1007/s00586-012-2381-1.
50. Weeks WB, Paraponaris A, Ventelou B. Geographic variation in rates of common surgical procedures in France in 2008-2010, and comparison to the US and Britain. *Health Policy*. 2014;118:215-221. doi:10.1016/j.healthpol.2014.08.015.
51. Walsh ME, Boland F, O'Byrne JM, Fahey T. Geographical variation in musculoskeletal surgical care in public hospitals in Ireland: a repeated cross-sectional study. *BMJ Open*. 2019;9:e028037. doi:10.1136/bmjopen-2018-028037.
52. Cherkin DC, Deyo RA, Loeser JD, Bush T, Waddell G. An international comparison of back surgery rates. *Spine*. 1994;19:1201-1206. doi:10.1097/00007632-199405310-00001.
53. Harris IA, Dao ATT. Trends of spinal fusion surgery in Australia: 1997 to 2006. *ANZ J Surg*. 2009;79:783-788. doi:10.1111/j.1445-2197.2009.05095.x.
54. Lopez CD, Boddapati V, Lombardi JM, et al. Medicare utilization and reimbursement for vertebroplasty and kyphoplasty. *Spine*. 2020;45:1744-1750. doi:10.1097/brs.0000000000003692.
55. Loeser JD, Konkelenberg RV, Volinn E, Cousins MJ. Small area analysis of lumbar spine surgery in South Australia. *ANZ J Surg*. 1993;63:14-19. doi:10.1111/j.1445-2197.1993.tb00026.x.
56. Yoshihara H, Yoneoka D. National trends in the surgical treatment for lumbar degenerative disc disease: United States, 2000 to 2009. *Spine J*. 2015;15:265-271. doi:10.1016/j.spinee.2014.09.026.
57. Peul WC, van Houwelingen HC, van den Hout WB, et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med*. 2007;356:2245-2256. doi:10.1056/NEJMoa064039.
58. Bailey CS, Rasoulinejad P, Taylor D, et al. Surgery versus conservative care for persistent sciatica lasting 4 to 12 months. *N Engl J Med*. 2020;382:1093-1102. doi:10.1056/NEJMoa1912658.
59. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation. *J Am Med Assoc*. 2006;296:2441-2450. doi:10.1001/jama.296.20.2441.
60. Hout WBv. d., Peul WC, Koes BW, Brand R, Kievit J, Thomeer RTWM. 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;1351-1354. doi;10.1136/bmj.39583.709074.
61. Volinn E, Mayer J, Diehr P, Koevering DV, Connell FA, Loeser JD. Small area analysis of surgery for low-back pain. *Spine*. 1992;17:575-581. doi:10.1097/00007632-199205000-00017.
62. Faciszewski T. Administrative databases in spine research. *Spine*. 1997;22:1270-1275.
63. Wennberg J. Dartmouth Atlas Projects: The Dartmouth Institute for Health Policy and Clinical Practice. Available from: https://www.dartmouthatlas.org/