Annette Kettler
Hans-Joachim Wilke

Review of existing grading systems for cervical or lumbar disc and facet joint degeneration

Received: 13 October 2004
Revised: 4 April 2005
Accepted: 14 April 2005
Published online: 20 September 2005
© Springer-Verlag 2005

Abstract The aim of this literature review was to present and to evaluate all grading systems for cervical and lumbar disc and facet joint degeneration, which are accessible from the MEDLINE database. A MEDLINE search was conducted to select all articles presenting own grading systems for cervical or lumbar disc or facet joint degeneration. To give an overview, these grading systems were listed systematically depending on the spinal region they refer to and the methodology used for grading. All systems were checked for reliability tests and those recommended for use having an interobserver Kappa or Intraclass Correlation Coefficient > 0.60 if disc degeneration was graded and > 0.40 if facet joint degeneration was graded. MEDLINE search revealed 42 different grading systems. Thirty of these were used to grade lumbar spine degeneration, ten were used to grade cervical spine degeneration and two were used to grade both. Thus, the grading systems for the lumbar spine represented the vast majority of all 42 grading systems. Interobserver reliability tests were found for 12 grading systems. Based

on their Kappa or Intraclass Correlation Coefficients nine of these could be recommended for use and three could not. All other systems could neither be recommended nor not be recommended since reliability tests were missing. These systems should therefore first be tested before use. The design of the grading systems varied considerably. Five grading systems were beginning with the lowest degree of degeneration, 37, however, with the normal, not degenerated state. A 5-grade scale was used in six systems, a 4-grade scale in 24, a 3-grade scale in eight and a 2-grade scale in three systems. In 15 cases the normal, not degenerated state was assigned to “grade 0”, in another 15 cases, however, this state was assigned to “grade 1”. This wide variety in the design of the grading systems makes comparisons difficult and may easily lead to confusion. We would therefore recommend to define certain standards. Our suggestion would be to use a scale of three to five grades, to begin the scale with the not degenerated state and to assign this state to “grade 0”.

A. Kettler · H.-J. Wilke (✉)
Institute of Orthopaedic Research and
Biomechanics, University of Ulm,
Ulm, Germany
E-mail:
hans-joachim.wilke@medizin.uni-ulm.de
Tel.: +49-731-50023482
Fax: +49-731-50023498

Purpose

Degeneration of the spine is a very common phenomenon. The morphological changes occurring during degeneration

have often been described macroscopically [13, 22, 24, 43, 60, 64, 66], histologically [6, 13, 16–18, 20, 22, 33, 64, 66], and using many different imaging techniques such as plain radiography [13, 24, 35, 36, 54], discography [1],

computed tomography [11, 22, 54, 71], or magnetic resonance imaging [22, 54, 73]. If these techniques have to be compared with each other in their ability to depict degeneration, grading systems are needed. Degeneration also has to be graded to investigate the relationship between morphology and functional parameters such as instability or intradiscal pressure, or to assess the correlation between different morphological changes. To answer whether disc degeneration causes facet joint degeneration also requires grading systems. Thus, grading systems for disc and facet joint degeneration are useful or even needed to answer a wide variety of different questions. It is therefore not surprising that more and more grading systems are reported in the literature.

The aim of this literature review was to collect, sort and evaluate all grading systems for cervical and lumbar disc and facet joint degeneration, which are accessible from the MEDLINE database.

Materials and methods

Literature search

A MEDLINE search was conducted via PubMed using various combinations of the following key words: degeneration, ageing, classification, grade, cervical spine, lumbar spine, intervertebral disc, and facet joint. Once a new grading system was found, the “related articles” option of PubMed was used to further expand the search. All extracted articles were reviewed and those selected with own grading systems for cervical or lumbar disc or facet joint degeneration. Even though this search was carried out thoroughly, it may be not complete.

According to the grading method, the selected articles were then assigned to one of the following subgroups: macroscopic anatomy, histology, plain radiography, computed tomography, magnetic resonance imaging, or discography if disc degeneration was graded, or to macroscopic anatomy, histology, plain radiography, conventional tomography, computed tomography, or magnetic resonance imaging if facet joint degeneration was graded.

Evaluation of grading systems

The evaluation of the grading systems was divided into two steps. In the first step, they were checked for reliability tests. Focus was put on the interobserver reliability, which is generally somewhat lower than the intraobserver reliability. To assess the reliability either Kappa [12] or Intraclass Correlation Coefficients (ICC) [57] have to be determined. According to the criteria published by Landis and Koch [37], a Kappa of <0.00

can be interpreted as poor agreement, 0.00–0.20 as slight agreement, 0.21–0.40 as fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial agreement, and >0.81 as almost perfect agreement. Based on this classification, we defined own criteria for recommendation: a grading system for cervical or lumbar disc degeneration was recommended only if Kappa or ICC was >0.60 (at least substantial agreement). Since the assessment of facet joint degeneration was assumed to be more difficult than that of the disc, the facet joint grading systems were recommended already at Kappa or ICC >0.40 (at least moderate agreement). Articles describing linear correlation coefficients instead of Kappa or ICC did not meet our criteria for recommendation since such correlation coefficients only reflect how strong the linear relationship but not how strong the agreement between two measurements is. If, for example, observer 1 would always grade half as high as observer 2, the linear correlation would be perfect but the agreement very low.

In the second step, all statistically recommended grading systems were additionally evaluated with respect to their clinical feasibility, clinical relevance and academic value.

Results

MEDLINE search revealed 32 articles describing 42 different grading systems for cervical or lumbar disc or facet joint degeneration. Thirty of these were applied to the lumbar spine, ten to the cervical spine and two to both spinal regions (Tables 1, 3, 5, 7).

The design of the grading systems varied considerably:

- *Beginning of the scale*: 37 grading systems were beginning with the normal, not degenerated state [1–3, 6, 10, 19, 26, 28–31, 35, 36, 38, 39, 47, 48, 52, 53, 55, 56, 58, 61–63, 67, 70, 74] while five grading systems were starting with the lowest degree of degeneration [8, 14, 34, 44].
- *Nomenclature of grades*: In 35 grading systems numerical grades such as grade 0, 1, 2, ... or category I, II, III ... [2, 3, 6, 8, 14, 19, 26, 28–31, 34–36, 38, 47, 48, 52, 53, 55, 58, 61–63, 67, 70, 74] and in seven systems terms such as “normal”, “mild”, “moderate”, or “severe” were used [1, 10, 39, 44, 56].
- *Nomenclature for the normal, not degenerated state*: If the scale was beginning with the normal, not degenerated state, this state was assigned to “grade 0” (15 grading systems [19, 31, 35, 36, 38, 52, 55, 58, 67, 70]), to “grade 1” (15 grading systems [2, 3, 26, 28–30, 34, 47, 48, 53, 61–63, 74]), or to a linguistic term such as “normal” or “not degenerated” (six grading systems: [10, 39, 56]).
- *Number of grades*: A 5-grade scale was used in six systems [1, 35, 36, 53, 58, 63], a 4-grade scale in 24 [2,

- 3, 19, 26, 28–31, 34, 44, 47, 48, 52, 56, 58, 67, 70], a 3-grade scale in eight [8, 38, 39, 55, 61, 62, 74] and a 2-grade scale in three systems [10, 14]. A score instead of distinct grades was used in one grading system [6].
- *Structure*: In 25 grading systems a certain combination of several morphological changes was assigned to each grade [2, 10, 14, 19, 26, 28, 31, 34, 44, 48, 52, 55, 56, 58, 61, 62, 67, 70, 74], in four others this list was divided into several subgroups such as “annulus”, “nucleus”, “endplates” for the disc or “cartilage”, “subchondral bone”, “margins” for the facet joints [3, 29, 53, 63]. In seven grading systems each feature of degeneration was graded separately and either summed up to a summary grade [38, 47] or not [8, 30, 39, 67].

Grading systems for lumbar disc degeneration

Twenty-two different grading systems for lumbar disc degeneration were found (Table 1), including five for

macroscopic anatomy, three for histology, six for plain radiography, five for magnetic resonance imaging, and three for discography. A grading system based on the computed tomography scans was not found probably because this imaging technique is not suitable to depict changes within the disc.

Interobserver reliability tests were found only for five grading systems (Table 2). Out of these five systems only those described by Thompson et al. [63] (Table 10), Boos et al. [6] (except for endplate variable “cells”) (Table 11), Lane et al. [38] (except for sclerosis) (Table 13), and Pfirrmann et al. [53] (Table 14) had substantial to almost perfect agreement and thus fulfilled our criteria for recommendation. Not recommended is the grading system of Madan et al. [44] since the most interobserver Kappa Coefficients were below 0.60. The reliability of the grading system of Gries et al. [29] and Berlemann et al. [3] could not be evaluated since Spearman’s correlation coefficients were determined instead of Kappa or Intraclass Correlation Coefficients [3, 29].

Table 1 Publications with own grading systems for lumbar disc degeneration (see Results). Chronological order within each subchapter

Lumbar disc degeneration			
Method	Grading system	Comments	See
Macroscopic anatomy	Nachemson 1960	Transverse sections	Table 9
	Galante 1967	Transverse sections	[26]
	Thompson et al. 1990 ^{a,b}	Sagittal sections	Table 10
Histology	Ziv et al. 1993	(Probably sagittal sections)	[74]
	Adams et al. 1996	Sagittal sections	[2]
	Gunzburg et al. 1992	Sagittal paraffin sections; HE staining	[31]
	Berlemann et al. 1998 and Gries et al. 2000	Sagittal paraffin sections; HE staining	[3, 29]
	Boos et al. 2002 ^{a,b}	Sagittal paraffin sections; HE, Masson-Goldner and Alcian blue-PAS staining	Table 11
Plain radiography	Kellgren and Lawrence 1952	Lateral views in flexion and extension and antero- posterior views	[35]
	Kellgren et al. 1963	Lateral views	[34]
	Gordon 1991	(Probably lateral and anteroposterior views)	[28]
	Mimura 1994	Lateral and antero-posterior views	Table 12
	Lane et al. 1993 ^{a,b}	Lateral views	Table 13
Magnetic resonance imaging	Madan et al. 2003 ^a	Lateral and antero-posterior views	[44]
	Schneiderman et al. 1987	T1-(disc height) and T2-weighted sagittal scans	[56]
	Butler et al. 1990	T2-weighted sagittal scans	[10]
	Tertti et al. 1991	T2-weighted sagittal scans	[62]
	Gunzburg et al. 1992	T2-weighted sagittal scans	[31]
Discography	Pfirrmann et al. 2001 ^{a,b}	T2-weighted sagittal scans	Table 14
	Adams et al. 1986	Lateral and some antero-posterior views	Table 15
	Schneiderman et al. 1987	Lateral and anteroposterior views (and axial CT scans)	[56]
	Gunzburg et al. 1992	Lateral and anteroposterior views	[31]

HE haematoxylin and eosin

^aGrading systems, which have been tested for reliability using Kappa or Intraclass Correlation Coefficients

^bRecommended grading systems (Kappa or ICC > 0.60, see Materials and methods)

Table 2 Reliability of grading systems for lumbar disc degeneration assessed in terms of Kappa or Intraclass Correlation Coefficients. Same order as in Table 1

Lumbar disc degeneration		
Grading system	Intraobserver reliability	Interobserver reliability
Thompson et al. (1990) ^a Table 10	Kappa [12]: 0.87–0.91 disc grades	Kappa [12] (three experienced observers): 0.69–0.94 disc grades
Boos et al. (2002) ^a (recommended for all variables except for endplate variable “cells”) Table 11	Not determined	Weighted Kappa (two observers involved in development of grading scheme): 0.786 (CI 0.726–0.846)–0.955 (CI 0.931–0.978) disc variables 0.787 (CI 0.705–0.870)–0.939 (CI 0.910–0.968) endplate variables Weighted Kappa (two observers not involved in development of grading scheme): 0.726 (CI 0.642–0.810)–0.806 (CI 0.727–0.885) disc variables 0.699 (CI 0.597–0.800)–0.977 (CI 0.947–1.009) endplate variables 0.493 (CI 0.398–0.588) endplate variable “cells”
Lane et al. (1993) ^a (recommended for narrowing, osteophytes, and summary grades only) Table 13	Intraclass Correlation Coefficient/Kappa [#] : 0.92 narrowing 0.96 osteophytes 0.59 sclerosis [#] 0.90 summary grades	Intraclass Correlation Coefficient/Kappa [#] (three experienced observers): 0.95 narrowing 0.91 osteophytes 0.55 sclerosis [#] 0.93 summary grades
Madan et al. (2003)	Not determined	Weighted Kappa (five observers with different levels of experience): 0.351–0.673 disc grades
Pfrrmann et al. (2001) ^a Table 14	Kappa [12]: 0.84–0.90 disc grades	Kappa [12] (three observers with different levels of experience): 0.74–0.81 disc grades

CI 95% confidence interval

^aRecommended grading systems (Kappa or ICC > 0.60, see [Materials and methods](#))

The clinical feasibility of the four statistically recommended grading systems of Thompson et al. [63] (Table 10), Boos et al. [6] (Table 11), Lane et al. [38] (Table 13) and Pfrrmann et al. [53] (Table 14) varies considerably: it is the highest for the grading system of Lane et al. since the only material needed are lateral radiographs and the system itself has a clear structure and can very easily be applied. The grading system of Pfrrmann et al. also has a high clinical feasibility, however, magnetic resonance images are needed and some more time is needed to arrive at the result. The clinical relevance, however, is higher for the grading system of Pfrrmann et al. than for that of Lane et al. since, on MRI images, the disc itself can directly be assessed, whereas on radiographs, only bony structures are assessable. The grading systems of Thompson et al. [63] and Boos et al. [6], in contrast, have more academic than clinical value since both systems can not be applied on patients. However, both of them are based on the detailed morphological studies, allow an in depth evaluation of the disc and are therefore valuable tools to grade disc degeneration in vitro.

Besides these four statistically recommended grading systems, another three are of special interest: the macroscopic grading system of Nachemson [48] (Table 9), the radiographic system of Mimura et al. [47] (Table 12)

and the discographic system of Adams et al. [1] (Table 15). Nachemson was one of the first to grade disc degeneration macroscopically. His grading system has often been used in the past. Mimura et al. were the first who tried to develop a more objective grading system. It is more based on numbers such as the percentage decrease in disc height than on subjective terms such as “mild”, “moderate”, or “severe”. Adams et al. finally need to be mentioned since they count among the first to describe the relationship between the distribution of fluid injected into the disc and disc degeneration.

Grading systems for cervical disc degeneration

To grade cervical disc degeneration ten systems were found (Table 3), two based on macroscopic anatomy, four using plain radiography, two using magnetic resonance imaging and two using discography. No histological or computed tomography grading systems were found.

Interobserver reliability tests were found only for the grading system of Kellgren et al. [34] (Tables 4, 16). Since the Kappa Coefficient was 0.71, this grading system fulfilled our criteria for recommendation.

From a clinical point of view this system has a good feasibility since it is based on the lateral radiographs and

Table 3 Publications with own grading systems for cervical disc degeneration (see [Results](#)). Chronological order within each subchapter

Cervical disc degeneration			
Method	Grading system	Comments	See
Macroscopic anatomy	Silberstein (1965)	Inspection of surface, no sections	[58]
	Viikari–Juntura et al. (1989)	Mid-frontal sections	[67]
Histology	No grading systems found		
Plain radiography	Kellgren and Lawrence (1952)	Lateral views in flexion and extension	[35]
	Kellgren et al. (1963) ^{a,b} (tested for reliability by Cote et al. 1997)	Lateral views	Table 16
	Brooker and Barter (1965)	Antero-posterior views, lateral views in neutral position, flexion and extension, and both oblique views	[8]
	Silberstein (1965)	Antero-posterior, lateral and both oblique views	[58]
	Viikari–Juntura et al. (1989)	Sagittal scans	[67]
Magnetic resonance imaging	Lehto et al. (1994)	T2-weighted scans	[39]
Discography	Viikari–Juntura et al. (1989)	Poster-anterior, lateral and axial views	[67]
	Schellhas et al. (1996)	Antero-posterior and lateral views	[55]

HE haematoxylin and eosin

^aGrading systems, which have been tested for reliability using Kappa or Intraclass Correlation Coefficients

^bRecommended grading systems (Kappa or ICC > 0.60, see [Materials and methods](#))

can easily be applied. In contrast, its clinical relevance is somewhat lower since the discs can only indirectly be assessed in terms of reactive changes of the surrounding bone and a loss of disc space height.

Grading systems for lumbar facet joint degeneration

Twelve grading systems for lumbar facet joint degeneration were found via MEDLINE (Table 5). These systems were either based on the macroscopic anatomy, histology, plain radiography, conventional tomography, computed tomography, or magnetic resonance imaging.

Interobserver reliability tests were found for five of these twelve grading systems (Table 6). As expected, the Kappa or Intraclass Correlation Coefficients were generally lower for the facet joints when compared to the intervertebral discs. Thus, only three grading systems fulfilled the criteria for recommendation, even though the threshold was lower than for the discs: Pathria et al. [52] (computed tomography only) (Table 17) and Weishaupt et al. [70] (computed tomography and magnetic resonance imaging) (Table 18). The Kappa Coefficients of the grading systems described by Pathria et al. [52] for plain radiography, and Coste et al. [14] were below 0.40 and should therefore not be used. The reliability of the grading systems of Gries et al. [29] and Kellgren and Lawrence [36] could not be assessed since linear correlation coefficients were reported instead of Kappa or ICC.

From a clinical point of view the statistically recommended grading systems of Pathria et al. [52] (Table 17) and Weishaupt et al. [70] (Table 18) are both feasible and relevant since they are based CT-scans or magnetic resonance images. In contrast, from an academic point of view, systems such as those of Tanno et al. [61] or Gries et al. [29] could be of interest since they allow to assess facet joint degeneration in vitro either by macroscopic or histologic inspection. However, these systems have not yet been tested for reliability.

Table 4 Reliability of grading systems for cervical disc degeneration assessed in terms of Kappa or Intraclass Correlation Coefficients

Cervical disc degeneration		
Grading system	Intraobserver reliability	Interobserver reliability
Kellgren et al. (1963) ^a (tested for reliability by Cote et al. 1997) Table 16 (and 19)	Not determined	Intraclass Correlation Coefficient [21, 59] (three experienced observers): 0.71 (CI 0.23–0.88) disc grades (0.45 (CI 0.09–0.71) facet joint grades)

CI 95% confidence interval

^aRecommended grading systems (Kappa or ICC > 0.60, see [Materials and methods](#))

Table 5 Publications with own grading systems for lumbar facet joint degeneration (see [Results](#)). Chronological order within each subchapter

Lumbar facet joint degeneration			
Method	Grading system	Comments	See
Macroscopic anatomy	Ziv et al. (1993)	Inspection of joint surface after India ink staining	[74]
	Tanno et al. (2003)	Inspection of joint surface	[61]
Histology	Gries et al. (2000)	Paraffin sections perpendicular to the joint line; HE staining	[29]
Plain radiography	Kellgren and Lawrence (1957)	Lateral views	[36]
Conventional tomography Computed tomography	Pathria et al. (1987) ^a	Both oblique views	[52]
	Demaerel et al. (1992)	Both oblique views	
	Pathria et al. (1987) ^{a,b}	Axial scans (0° gantry angulation)	Table 17
	Butler et al. (1990)	(probably axial scans)	[10]
	Coste et al. (1994) ^a	Axial scans (parallel to disc space)	[14]
Magnetic resonance imaging	Weishaupt et al. (1999) ^{a,b}	Axial scans	Table 18
	Grogan et al. (1997)	T1- and T2-weighted axial scans	[30]
	Weishaupt et al. (1999) ^{a,b}	T1- and T2-weighted sagittal, and T2-weighted axial scans	Table 18

HE haematoxylin and eosin

^aGrading systems, which have been tested for reliability using Kappa or Intraclass Correlation Coefficients

^bRecommended grading systems (Kappa or ICC > 0.40, see [Materials and methods](#))

Grading systems for cervical facet joint degeneration based on the macroscopic anatomy, and three on plain radiography. No histologic grading system and no systems using conventional tomography, computed tomography or magnetic resonance imaging were found.

Cervical facet joint degeneration was only graded in four articles (Table 7). One of these four grading systems was

Table 6 Reliability of grading systems for lumbar facet joint degeneration assessed in terms of Kappa or Intraclass Correlation Coefficients. Same order as in Table 5

Lumbar facet joint degeneration		
Grading system	Intraobserver reliability	Interobserver reliability
Pathria et al. (1987) ^a (recommended for CT only) Table 17	Not determined	Kappa (two experienced observers): 0.46 facet joint grades on CT 0.26 facet joint grades on oblique radiographs
Coste et al. (1994)	Unweighted Kappa [21]: −0.03–0.34 right facet joints at L4-5 level 0.30–0.74 right facet joints at L5-S1 level −0.06–0.48 left facet joints at L4-5 level 0.26–0.68 left facet joints at L5-S1 level	Unweighted Kappa [21] (four experienced observers): 0.02–0.09 right facet joints at L4-5 level 0.11–0.39 right facet joints at L5-S1 level −0.11–0.10 left facet joints at L4-5 level 0.10–0.39 left facet joints at L5-S1 level
Weishaupt et al. (1999) ^a Table 18	Weighted Kappa [37]: 0.70 facet joint grades on MRI, observer 1 0.76 facet joint grades on MRI, observer 2 0.70 facet joint grades on CT, observer 1 0.77 facet joint grades on CT, observer 2	Weighted Kappa [37] (two experienced observers): 0.41 facet joint grades on MRI 0.60 facet joint grades on CT

CI 95% Confidence interval, MRI magnetic resonance imaging, CT computed tomography

^aRecommended grading systems (Kappa or ICC > 0.40, see [Materials and methods](#))

Table 7 Publications with own grading systems for cervical facet joint degeneration (see [Results](#)). Chronological order within each subchapter

Cervical facet joint degeneration			
Method	Grading system	Comments	See
Macroscopic anatomy	Silberstein (1965)	Probably inspection of surface	[58]
Histology	No grading systems found		
Plain radiography	Kellgren and Lawrence (1957)	Lateral views	[36]
	Kellgren et al. (1963) ^{a,b} (tested for reliability by Cote et al. 1997)	Lateral views	Table 19
	Silberstein (1965)	Anteroposterior, lateral and both oblique views	[58]
Conventional tomography	No grading systems found		
Computed tomography	No grading systems found		
Magnetic resonance imaging	No grading systems found		

^aGrading systems, which have been tested for reliability using Kappa or Intraclass Correlation Coefficients

^bRecommended grading systems (Kappa or ICC > 0.40, see [Materials and methods](#))

The interobserver reliability of the grading system of Kellgren et al. [34] (Table 19) was only one fulfilling the criteria for recommendation (Table 8). The system described by Kellgren and Lawrence [36] was only tested for linear correlation. Thus, its reliability still needs to be determined.

From a clinical point of view the recommended grading system of Kellgren et al. [34] is feasible since only lateral radiographs are needed.

Discussion

Forty-two different grading systems for cervical and lumbar disc and facet joint degeneration were accessible from the MEDLINE database. This high number is not astonishing since grading systems are needed to answer a wide variety of different questions. Several grading systems were for example used to investigate the correlation between morphology and flexibility, intradiscal pressure, or pain [8, 19, 25, 31, 35, 47, 48, 55]. Others were needed to compare different diagnostic methods with each other [1, 3, 27, 52, 55, 56, 62, 67, 70], or to investigate the relation between the disc and facet joint degeneration [10, 29, 61, 74].

The majority of all 42 grading systems were used for the lumbar but only few for the cervical spine. Histologic grading systems, for example, were only found for the lumbar but not for the cervical spine. This also applies to grading systems for facet joint degeneration using the conventional tomography, computed tomography or magnetic resonance imaging. Thus, despite of their high number, new grading systems may still be needed especially for the cervical spine.

Regardless of the spinal region the grading systems were belonging to, the variability of their design was very high: Some systems started with the normal, not degenerated state, others with the lowest degree of degeneration; some assigned the not degenerated state to “grade 0”, others to “grade 1”; some were using a 5-grade scale, others a 4-grade, 3-grade, or 2-grade scale; some assigned linguistic terms such as “normal”, “mild”, “severe” to each grade, some others numerical values (grade 0, grade 1, grade 2, ...). These and several other variations make comparisons difficult and may easily lead to confusion. We would therefore recommend to define certain standards. Whether to start with “grade 0” or “grade 1” or with the not degenerated or the least degenerated state and whether to use linguistic terms or numerical values to grade degeneration could easily be standardised. Our suggestion would be to

Table 8 Reliability of grading systems for cervical facet joint degeneration assessed in terms of Kappa or Intraclass Correlation Coefficients

Cervical facet joint degeneration		
Grading system	Intraobserver reliability	Interobserver reliability
Kellgren et al. 1963 ^a (tested for reliability by Cote et al. 1997) Table (16 and) 19	Not determined	Intraclass correlation coefficient [21, 59] (three experienced observers): (0.71 (CI 0.23–0.88) disc grades) 0.45 (CI 0.09–0.71) facet joint grades

CI 95% Confidence interval

^aRecommended grading systems (Kappa or ICC > 0.40, see [Materials and methods](#))

begin the scale with the not degenerated state and to assign this state to “grade 0”. This is what many authors already did. For the number of grades, however, it is critical to define a standard since this number should individually be adapted to the purpose of the study and the number of discs or facet joints, which have to be graded. A small number of discs or facet joints should be divided into less degrees of degeneration than a large number since otherwise, the number of individual discs or facet joints in each group of degeneration may become too small for further evaluation. Furthermore, the method used for grading also needs to be considered since most methods have a limited resolution and only cover certain but not all structures of the discs or facet joints. Thus, if the number of discs or facet joints is small and the method does not allow a detailed inspection of all structures such as plain radiography, the use of only three or four degrees of degeneration can be recommended. If, in contrast, the number of discs or facet joints is high and methods such as histology or macroscopy are used, which allow an in depth inspection of all tissues, up to five degrees of degeneration can be recommended. However, in general less than three grades should not be used since otherwise discs or facet joints with a wide variety of different appearances would all be mixed together within one and the same degree of degeneration. Sometimes it might even be advantageous not to report an overall degree of degeneration but rather the appearance of the different structures individually, because otherwise details concerning these individual structures might get lost.

Reliability is one of the most important characteristics of a grading system. Intra- and interobserver reliability can be measured in terms of Kappa [12] or ICC [57]. Unfortunately, only few of all reviewed grading systems were tested for reliability. Out of these, nine grading systems fulfilled our criteria for recommendation (Kappa or ICC > 0.60 for discs and > 0.40 for facet joints) but one for lumbar disc and two for lumbar facet joint degeneration did not (Tables 2, 4, 6, 8) [14, 44, 52]. These results indicate that grading systems should always be tested for reliability before use. Thresholds for Kappa or ICC might help to decide whether to use or not to use a tested system. General thresholds such as ours are suitable, if, for example, a general comparison between several grading systems is needed. However, if a specific grading system is planned to be used, the thresholds need to be adjusted to the respective purpose of the study and to the degree of experience and training of the observers, who will use the system and who had used it to determine its reliability. The thresholds should also be modified if templates will be or were available and if the observers were involved in the development of the grading system.

Conclusions and recommendations

- Many different systems to grade lumbar disc or facet joint degeneration were accessible from the MEDLINE database. For the cervical spine, this number was much smaller.
- Missing standards concerning the design of the grading systems make comparisons difficult and may easily lead to confusion. We would therefore recommend to define certain standards. Our suggestion would be to use a scale of three to five grades, to begin the scale with the not degenerated state and to assign this state to “grade 0”.
- From a statistical point of view, the following nine of all 42 reviewed grading systems can be recommended for use since their Kappa or ICC for interobserver reliability were > 0.60 (discs) or > 0.40 (facet joints):
 1. Thompson et al. [63]: lumbar disc, macroscopic anatomy (Table 10)
 2. Boos et al. [6]: lumbar disc, histology (except for endplate variable “cells”) (Table 11)
 3. Lane et al. [38]: lumbar disc, plain radiography (except for sclerosis) (Table 13)
 4. Pfirrmann et al. [53]: lumbar disc, magnetic resonance imaging (Table 14)
 5. Kellgren et al. 1963: cervical disc, plain radiography (Table 16)
 6. Pathria et al. [52]: lumbar facet joints, computed tomography (Table 17)
 7. Weishaupt et al. [70]: lumbar facet joints, computed tomography (Table 18)
 8. Weishaupt et al. [70]: lumbar facet joints, magnetic resonance imaging (Table 18)
 9. Kellgren et al. 1963: cervical facet joints, plain radiography (Table 19)
- The interobserver reliability of the following three grading systems was ≤ 0.60 (discs) or ≤ 0.40 (facet joints). These systems are therefore not recommended for use:
 1. Madan et al. [44]: lumbar disc, plain radiography
 2. Pathria et al. [52]: lumbar facet joints, plain radiography
 3. Coste et al. [74]: lumbar facet joint degeneration, computed tomography
- For all other grading systems our criteria for recommendation were not applicable since they were not tested for reliability. Since these systems could be more or less reliable than those recommended above, they should at least be tested for interobserver agreement before use.

Acknowledgements The authors would specially like to thank the *Deutsche Arthrose-Hilfe e.V.* and the German Research Council (DFG, WI 1352/6-1) for financial support.

Appendix

Grading systems for lumbar disc degeneration

Macroscopic anatomy

Table 9 Macroscopic grading of lumbar disc degeneration on transverse sections described by Nachemson 1960 [48]

Group 1	Discs without changes visible to the naked eye. In these cases a gelatinous shiny nucleus pulposus was seen; it was easily delimited from the annulus fibrosus, which was free from macroscopic ruptures.
Group 2	Discs which showed macroscopic changes in the nucleus pulposus. The nucleus was somewhat more fibrous, but could be clearly distinguished from the annulus, which was intact.
Group 3	Specimens which showed macroscopic changes in both the nucleus pulposus and the annulus fibrosus. The nucleus in these discs was more fibrotic but still soft. The boundary between nucleus and annulus was no longer distinct, but could still be seen. Changes in the annulus fibrosus consisted of isolated fissures.
Group 4	Specimens which showed more severe macroscopic changes. The discs in this group exhibited fissure formation and cavities in both the nucleus and the annulus. Marginal osteophytes were often found in adjoining vertebrae.

Based on Frieberg and Hirsch (1949), Lindblom (1944, 1951), Virgin (1951, 1958), Naylor and Smare (1953), Hendry (1958) [24, 32, 40–42, 49, 68, 69]

The purpose was to investigate the effect of disc degeneration on intradiscal pressure

Table 10 Macroscopic grading of lumbar disc degeneration on sagittal sections proposed by Thompson et al. [63]

Grade	Nucleus	Annulus	Endplate	Vertebral body
I	Bulging gel	Discrete fibrous lamellas	Hyaline, uniformly thick	Margins rounded
II	White fibrous tissue peripherically	Mucinous material between lamellas	Thickness irregular	Margins pointed
III	Consolidated fibrous tissue	Extensive mucinous infiltration; loss of annular-nuclear demarcation	Focal defects in cartilage	Early chondrophytes or osteophytes at margins
IV	Horizontal clefts parallel to endplate	Focal disruptions	Fibrocartilage extending from subchondral bone; irregularity and focal sclerosis in subchondral bone	Osteophytes less than 2 mm
V	Clefts extended through nucleus and annulus		Diffuse sclerosis	Osteophytes greater than 2 mm

Based on Vernon-Roberts 1987 [65]

The purpose was to provide an epidemiologically tested method of assessing the morphologic state of the disc. Tested for comprehensiveness, intra- and interobserver agreement, ability to discriminate between grades, and credibility

Histology

Table 11 Histologic grading of lumbar disc degeneration on sagittal paraffin sections stained with haematoxylin and eosin, Masson–Goldner and Alcian blue-PAS proposed by Boos et al. [6]

Intervertebral Disc	Endplate
<p>Cells (chondrocyte proliferation)</p> <p>0 = No proliferation 1 = Increased cell density 2 = Connection of two chondrocytes 3 = Small size clones (i.e. several chondrocytes group together, i.e., 2–7 cells) 4 = Moderate size clones (i.e., > 8 cells) 5 = Huge clones (i.e., 15 cells) 6 = Scar/tissue defects</p> <p>Mucous degeneration</p> <p>0 = Absent 1 = Rarely present 2 = Present in intermediate amounts 3 = Abundantly present 4 = Scar/tissue defects</p> <p>Cell death</p> <p>0 = Absent 1 = Rarely present 2 = Present in intermediate amounts 3 = Abundantly present 4 = Scar/tissue defects</p> <p>Tear and cleft formation</p> <p>0 = Absent 1 = Rarely present 2 = Present in intermediate amounts 3 = Abundantly present 4 = Scar/tissue defects</p> <p>Granular changes</p> <p>0 = Absent 1 = Rarely present 2 = Present in intermediate amounts 3 = Abundantly present 4 = Scar/tissue defects</p> <p>Maximum: 22 points</p>	<p>Cells</p> <p>0 = Normal cellularity 1 = Localized cell proliferation 2 = Diffuse cell proliferation 3 = Extensive cell proliferation 4 = Scar/tissue defects</p> <p>Cartilage disorganisation</p> <p>0 = Well-structured hyaline cartilage 1 = Cartilage irregularities (obliterated vessels?) 2 = Disorganised matrix with thinning 3 = Complete cartilage disorganisation with defects 4 = Scar/tissue defects</p> <p>Cartilage cracks</p> <p>0 = Absent 1 = Rarely present 2 = Present in intermediate amounts 3 = Abundantly present 4 = Scar/tissue defects</p> <p>Microfracture</p> <p>0 = Absent 1 = Present 2 = Scar/tissue defects</p> <p>New bone formation</p> <p>0 = Absent 1 = Present 2 = Scar/tissue defects</p> <p>Bony sclerosis</p> <p>0 = Absent 1 = Present 2 = Scar/tissue defects</p> <p>Maximum: 18 points</p>

Based on Boos et al., Buckwalter et al., Coventry et al., Fraser et al., Milgram, Nerlich et al., Osti et al., Yasuma et al. [4, 9, 16–18, 23, 46, 50, 51, 72] The purpose was to develop a practicable and reliable classification system for age-related histologic disc alteration. Tested for interobserver reliability, practicability, and validity

Plain radiography

Table 12 Radiographic grading of lumbar disc degeneration on lateral and antero-posterior views described by Mimura et al. [47]

Disc height changes (% of adjacent discs)	Osteophyte formation (Sum of points on 8 Edges: < 3 mm 1 pt., > 3 mm 2 points.)	Endplate sclerosis
0 = Normal 1 = Mild (> 75%) 2 = Moderate (> 50%) 3 = Severe (> 25%) 4 = Very severe (< 25%)	0 = 0 point 1 = 1–4 points 2 = 5–8 points 3 = 9–12 points 4 = 13–16 points	0 = None 1 = Either endplate 2 = Both endplates

The purpose was to investigate the effect of degeneration on flexibility. According to the sum of the disc height, osteophyte, and sclerosis scores, each disc was graded on a scale of 1 to 4 with: Grade 1 being 0.0 to 1.5 points
Grade 2 being 2.0 to 3.5 points
Grade 3 being 4.0 to 6.0 points
Grade 4 being > 6.0 points

Table 13 Radiographic grading of lumbar disc degeneration on lateral views described by Lane et al. 1993 [38]

	Joint space narrowing	Osteophytes anterior and posterior	Sclerosis
0	None	None	None
1	Definite (mild) narrowing	Small	Present
2	Moderate	Moderate	–
3	Severe (complete loss of joint space)	Large	–

The purpose was to develop a method for assessing the presence and severity of radiographic features of lumbar disc degeneration. Joint space narrowing, anterior and posterior osteophyte formation, and subchondral sclerosis were assessed individually. Based on these features, an overall grading was given from 0 to 2: Grade 0 = Normal joint (0 for osteophytes and narrowing)
Grade 1 = Mild (1) narrowing or mild (1) osteophytes
Grade 2 = Moderate-severe (2–3) narrowing and/or moderate-severe (2–3) osteophytes
Tested for intra- and interobserver reliability

Magnetic resonance imaging

Table 14 Grading of lumbar disc degeneration on T2-weighted sagittal MRI-scans proposed by Pfirrmann et al. [53]

Grade	Structure	Distinction of Nucleus and Annulus	Signal intensity	Height of Intervertebral Disc
I	Homogeneous, bright white	Clear	Hyperintense, isointense to cerebrospinal fluid	Normal
II	Inhomogeneous with or without horizontal bands	Clear	Hyperintense, isointense to cerebrospinal fluid	Normal
III	Inhomogeneous, grey	Unclear	Intermediate	Normal to slightly decreased
IV	Inhomogeneous, grey to black	Lost	Intermediate to hypointense	Normal to moderately decreased
V	Inhomogeneous, black	Lost	Hypointense	Collapsed disc space

Based on Boos et al., Brant-Zawadzki et al., Masaryk et al., Thompson et al. [5, 7, 45, 63]

The purpose was to develop an applicable, reliable classification system. Tested for intra- and interobserver reliability and ability to distinct between grades

Discography

Table 15 Discographic grading of lumbar disc degeneration in the sagittal plane proposed by Adams et al. [1]

Discogram type	Stage of disc degeneration
1. Cottonball	No signs of degeneration.
2. Lobular	Soft, white amorphous nucleus
3. Irregular	Mature disc with nucleus starting to coalesce into fibrous lumps
4. Fissured	Degenerated disc with fissures and clefts in the nucleus and inner annulus
5. Rupture	Degenerated disc with radial fissure leading to the outer edge of the annulus
	Disc has a complete radial fissure that allows injected fluid to escape. Can be in any state of degeneration

The purpose was to depict the stages of disc degeneration by means of discograms

Grading systems for cervical disc degeneration

Plain radiography

Table 16 Radiographic grading of cervical disc degeneration on lateral views described by Kellgren et al. [34]

Grade 1	Minimal anterior osteophytosis
Grade 2	Definite anterior osteophytosis with possible narrowing of the disc space and some sclerosis of vertebral plates
Grade 3	Moderate narrowing of the disc space with definite sclerosis of vertebral plates and osteophytosis
Grade 4	Severe narrowing of the disc space with sclerosis of vertebral plates and multiple large osteophytes

The purpose was to present some agreed system of grading radiological changes in the joints. Tested for interobserver reliability by Coté et al. [15], who added an additional Grade 0 = Absence of degeneration in the disc

Grading systems for lumbar facet joint degeneration

Computed tomography

Table 17 Grading of lumbar facet joint degeneration on oblique conventional radiographs and CT scans described by Pathria et al. [52]

Grade 0	Normal
Grade 1	Joint space narrowing (mild degenerative disease)
Grade 2	Narrowing plus sclerosis or hypertrophy (moderate degenerative disease)
Grade 3	Severe osteoarthritis with narrowing, sclerosis, and osteophytes (severe degenerative disease)

The purpose was to assess the accuracy of oblique radiographs in the detection of lumbar facet joint osteoarthritis. Tested for interobserver reliability

Magnetic resonance imaging

Table 18 Grading of lumbar facet joint degeneration using computed tomography and magnetic resonance imaging described by Weishaupt et al. [70]

Grade 0	Normal facet joint space (2–4 mm width)
Grade 1	Narrowing of the facet joint space (< 2 mm) and/or small osteophytes and/or mild hypertrophy of the articular process
Grade 2	Narrowing of the facet joint space and/or moderate osteophytes and/or moderate hypertrophy of the articular process and/or mild subarticular bone erosions
Grade 3	Narrowing of the facet joint space and/or large osteophytes and/or severe hypertrophy of the articular process and/or severe subarticular bone erosions and/or subchondral cysts

Based on Pathria et al. [52]

The purpose was to test the agreement between MR imaging and CT in the assessment of osteoarthritis of the lumbar facet joints. Tested for inter- and intraobserver reliability

Grading systems for cervical facet joint degeneration

Plain radiography

Table 19 Radiographic grading of cervical facet joint degeneration on lateral views described by Kellgren et al. [34]

Grade 1	Doubtful osteophytes on margins of the articular facets of apophyseal joints
Grade 2	Definite osteophytes and subchondral sclerosis in apophyseal joints
Grade 3	Moderate osteophytes, subchondral sclerosis and some irregularity of articular facets
Grade 4	Many large osteophytes and severe sclerosis and irregularity of the apophyseal joints

The purpose was to present some agreed system of grading radiological changes in the joints. Tested for interobserver reliability by Coté et al. [15], who added an additional Grade 0 = Absence of degeneration in the apophyseal joints

References

- Adams MA, Dolan P, Hutton WC (1986) The stages of disc degeneration as revealed by discograms. *J Bone Joint Surg Br* 68(1):36–41
- Adams MA, McNally DS, Dolan P (1996) 'Stress' distributions inside intervertebral discs. The effects of age and degeneration. *J Bone Joint Surg Br* 78(6):965–972
- Berlemann U, Gries NC, Moore RJ (1998) The relationship between height, shape and histological changes in early degeneration of the lower lumbar discs. *Eur Spine J* 7(3):212–217
- Boos N, Nerlich AG, Wiest I, von der Mark K, Aebi M (1997) Immunolocalization of type X collagen in human lumbar intervertebral discs during ageing and degeneration. *Histochem Cell Biol* 108(6):471–480
- Boos N, Wallin A, Schmucker T, Aebi M, Boesch C (1994) Quantitative MR imaging of lumbar intervertebral disc and vertebral bodies: methodology, reproducibility, and preliminary results. *Magn Reson Imaging* 12(4):577–587
- Boos N, Weissbach S, Rohrbach H, Weiler C, Spratt KF, Nerlich AG (2002) Classification of age-related changes in lumbar intervertebral discs: 2002 Volvo Award in basic science. *Spine* 27(23):2631–2644
- Brant-Zawadzki MN, Jensen MC, Obuchowski N, Ross JS, Modic MT (1995) Interobserver and intraobserver variability in interpretation of lumbar disc abnormalities. A comparison of two nomenclatures (discussion 1264). *Spine* 20(11):1257–1263
- Brooker AE, Barter RW (1965) Cervical spondylosis. A clinical study with comparative radiology. *Brain* 88(5):925–936
- Buckwalter JA (1995) Aging and degeneration of the human intervertebral disc. *Spine* 20(11):1307–1314
- Butler D, Trafimow JH, Andersson GB, McNeill TW, Huckman MS (1990) Discs degenerate before facets. *Spine* 15(2):111–113
- Carrera GF, Haughton VM, Syvertsen A, Williams AL (1980) Computed tomography of the lumbar facet joints. *Radiology* 134(1):145–148
- Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 70:213–220
- Collins DH (1949) The pathology of articular and spinal diseases. Edward Arnold & Co, London

14. Coste J, Judet O, Barre O, Siaud JR, Cohen de Lara A, Paolaggi JB (1994) Inter- and intraobserver variability in the interpretation of computed tomography of the lumbar spine. *J Clin Epidemiol* 47(4):375–381
15. Cote P, Cassidy JD, Yong-Hing K, Sibley J, Loewy J (1997) Apophysial joint degeneration, disc degeneration, and sagittal curve of the cervical spine. Can they be measured reliably on radiographs? *Spine* 22(8):859–864
16. Coventry MB (1945) The intervertebral disc: Its macroscopic anatomy and pathology: Part III. Pathological changes in the intervertebral disc lesion. *J Bone Joint Surg [Am]* 27:460–473
17. Coventry MB (1945) The intervertebral disc: Its microscopic anatomy and pathology: Part I. Anatomy, development and pathology. *J Bone Joint Surg [Am]* 27:105–112
18. Coventry MB (1945) The intervertebral disc: Its microscopic anatomy and pathology: Part II. Changes in the intervertebral disc concomitant with age. *J Bone Joint Surg [Am]* 27:233–247
19. Demaerel P, Wilms G, Goffin J, Baert AL (1992) Osteoarthritis of the facet joints and its role in low-back pain: evaluation with conventional tomography. *J Belge Radiol* 75(2):81–86
20. Eisenstein SM, Parry CR (1987) The lumbar facet arthrosis syndrome. Clinical presentation and articular surface changes. *J Bone Joint Surg Br* 69(1):3–7
21. Fleiss JL (1981) Statistical methods for rates and proportions, 2nd edn. Wiley, New York
22. Fletcher G, Haughton VM, Ho KC, Yu SW (1990) Age-related changes in the cervical facet joints: studies with cryomicrotomy, MR, and CT. *AJR Am J Roentgenol* 154(4):817–820
23. Fraser RD, Osti OL, Vernon-Roberts B (1993) Intervertebral disc degeneration. *Eur Spine J* 1:205–213
24. Friberg S, Hirsch C (1949) Anatomical and clinical studies on lumbar disc degeneration. *Acta Orthop Scand* 19:222–242
25. Friedenberg ZB, Miller WT (1963) Degenerative disc disease of the cervical spine. *J Bone Joint Surg Am* 45:1171–8
26. Galante JO (1967) Tensile properties of the human lumbar annulus fibrosus. *Acta Orthop Scand (Suppl100)*:1–91
27. Gibson MJ, Buckley J, Mawhinney R, Mulholland RC, Worthington BS (1986) Magnetic resonance imaging and discography in the diagnosis of disc degeneration. A comparative study of 50 discs. *J Bone Joint Surg Br* 68(3):369–373
28. Gordon SJ, Yang KH, Mayer PJ, Mace AH Jr, Kish VL, Radin EL (1991) Mechanism of disc rupture. A preliminary report. *Spine* 16(4):450–456
29. Gries NC, Berlemann U, Moore RJ, Vernon-Roberts B (2000) Early histologic changes in lower lumbar discs and facet joints and their correlation. *Eur Spine J* 9(1):23–29
30. Grogan J, Nowicki BH, Schmidt TA, Haughton VM (1997) Lumbar facet joint tropism does not accelerate degeneration of the facet joints. *AJNR Am J Neuroradiol* 18(7):1325–1329
31. Gunzburg R, Parkinson R, Moore R et al (1992) A cadaveric study comparing discography, magnetic resonance imaging, histology, and mechanical behavior of the human lumbar disc. *Spine* 17(4):417–426
32. Hendry NG (1958) The hydration of the nucleus pulposus and its relation to intervertebral disc derangement. *J Bone Joint Surg Br* 40B(1):132–144
33. Hirsch C (1972) Some morphological changes in the cervical spine during ageing. In: Hirsch C, Zotterman Y (eds) *Cervical pain*. Pergamon Press, New York, pp 21–32
34. Kellgren JH, Jeffrey MR, Ball J (1963) The epidemiology of chronic rheumatism. vol II: Atlas of standard radiographs of arthritis. Blackwell Scientific Publications, Oxford, pp 14–19
35. Kellgren JH, Lawrence JS (1952) Rheumatism in miners. II. X-ray study. *Br J Ind Med* 9(3):197–207
36. Kellgren JH, Lawrence JS (1957) Radiological assessment of osteoarthritis. *Ann Rheum Dis* 16(4):494–502
37. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159–174
38. Lane NE, Nevitt MC, Genant HK, Hochberg MC (1993) Reliability of new indices of radiographic osteoarthritis of the hand and hip and lumbar disc degeneration. *J Rheumatol* 20(11):1911–1918
39. Lehto JJ, Terviti MO, Komu ME, Pajanen HE, Tuominen J, Kormanen MJ (1994) Age-related MRI changes at 0.1 T in cervical discs in asymptomatic subjects. *Neuroradiology* 36(1):49–53
40. Lindblom K (1944) Protrusions of discs and nerve compression in the lumbar region. *Acta Radiol* 25:195
41. Lindblom K (1951) Backache and its relation to ruptures of the intravertebral disks. *Radiology* 57(5):710–719
42. Lindblom K (1951) Discography of dissecting transosseous ruptures of intervertebral discs in the lumbar region. *Acta Radiol* 36(1):12–16
43. Macnab I (1975) Cervical spondylosis. *Clin Orthop* (109):69–77
44. Madan SS, Rai A, Harley JM (2003) Interobserver error in interpretation of the radiographs for degeneration of the lumbar spine. *Iowa Orthop J* 23:51–56
45. Masaryk TJ, Ross JS, Modic MT, Boumphey F, Bohlman H, Wilber G (1988) High-resolution MR imaging of sequestered lumbar intervertebral disks. *AJR Am J Roentgenol* 150(5):1155–1162
46. Milgram J (1990) Intervertebral disc disease and degenerative arthritis of the spine. In: Milgram J (ed) *Radiologic and histologic pathology of nontumorous diseases of bone and joints*. Northbrook Publishing Company, pp 519–588
47. Mimura M, Panjabi MM, Oxland TR, Crisco JJ, Yamamoto I, Vasavada A (1994) Disc degeneration affects the multidirectional flexibility of the lumbar spine. *Spine* 19(12):1371–1380
48. Nachemson A (1960) Lumbar Intradiscal Pressure. Experimental studies on post-mortem material. *Acta Orthop Scand Suppl* 43:43
49. Naylor A, Smare DL (1953) Fluid content of the nucleus pulposus as a factor in the disk syndrome; preliminary report. *Br Med J* 4843:975–976
50. Nerlich AG, Schleicher ED, Boos N (1997) 1997 Volvo Award winner in basic science studies. Immunohistologic markers for age-related changes of human lumbar intervertebral discs. *Spine* 22(24):2781–2795
51. Osti OL, Vernon-Roberts B, Fraser RD (1990) 1990 Volvo Award in experimental studies. Annulus tears and intervertebral disc degeneration. An experimental study using an animal model. *Spine* 15(8):762–767
52. Pathria M, Sartoris DJ, Resnick D (1987) Osteoarthritis of the facet joints: accuracy of oblique radiographic assessment. *Radiology* 164(1):227–230
53. Pfirrmann CW, Metzger A, Zanetti M, Hodler J, Boos N (2001) Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* 26(17):1873–1878
54. Resnick D (1985) Degenerative diseases of the vertebral column. *Radiology* 156(1):3–14

55. Schellhas KP, Smith MD, Gundry CR, Pollei SR (1996) Cervical discogenic pain. Prospective correlation of magnetic resonance imaging and discography in asymptomatic subjects and pain sufferers (discussion 311–312). *Spine* 21(3):300–311
56. Schneiderman G, Flannigan B, Kingston S, Thomas J, Dillin WH, Watkins RG (1987) Magnetic resonance imaging in the diagnosis of disc degeneration: correlation with discography. *Spine* 12(3):276–281
57. Shrout PE, Fleiss J (1979) Intraclass correlations: Uses in Assessing Rater Reliability. *Psychol Bull* 86(2):420–428
58. Silberstein CE (1965) The Evolution of Degenerative Changes in the Cervical Spine and an Investigation into the “Joints of Luschka”. *Clin Orthop* 40:184–204
59. Streiner DL, Norman GL (1989) Health measurement scales: a practical guide to their development and use. Oxford University Press, Oxford
60. Swanepoel MW, Adams LM, Smeathers JE (1995) Human lumbar apophyseal joint damage and intervertebral disc degeneration. *Ann Rheum Dis* 54(3):182–188
61. Tanno I, Oguma H, Murakami G, Sato S, Yamashita T (2003) Which portion in a facet is specifically affected by articular cartilage degeneration with aging in the human lumbar zygapophysial joint? *Okajimas Folia Anat Jpn* 80(1):29–34
62. Terviti M, Paajanen H, Laato M, Aho H, Komu M, Kormanen M (1991) Disc degeneration in magnetic resonance imaging. A comparative biochemical, histologic, and radiologic study in cadaver spines. *Spine* 16(6):629–634
63. Thompson JP, Pearce RH, Schechter MT, Adams ME, Tsang IK, Bishop PB (1990) Preliminary evaluation of a scheme for grading the gross morphology of the human intervertebral disc. *Spine* 15(5):411–415
64. Töndury G (1972) The behaviour of the cervical discs during life. In: Hirsch C, Zotterman Y (eds) *Cervical pain*. Pergamon Press, New York, pp 59–66
65. Vernon-Roberts B (1987) Pathology of intervertebral discs and apophyseal joints. The lumbar spine and back pain. Churchill-Livingstone, London, pp 37–55
66. Vernon-Roberts B, Pirie CJ (1977) Degenerative changes in the intervertebral discs of the lumbar spine and their sequelae. *Rheumatol Rehabil* 16(1):13–21
67. Viikari-Juntura E, Raininko R, Videman T, Porkka L (1989) Evaluation of cervical disc degeneration with ultralow field MRI and discography. An experimental study on cadavers. *Spine* 14(6):616–619
68. Virgin WJ (1951) Experimental investigations into the physical properties of the intervertebral disc. *J Bone Joint Surg Br* 33B(4):607–611
69. Virgin WJ (1958) Anatomical and pathological aspects of the intervertebral disc. *Indian J Surg* 20:113
70. Weishaupt D, Zanetti M, Boos N, Hodler J (1999) MR imaging and CT in osteoarthritis of the lumbar facet joints. *Skeletal Radiol* 28(4):215–219
71. Wybier M (2001) Imaging of lumbar degenerative changes involving structures other than disk space. *Radiol Clin North Am* 39(1):101–114
72. Yasuma T, Koh S, Okamura T, Yamauchi Y (1990) Histological changes in aging lumbar intervertebral discs. Their role in protrusions and prolapses. *J Bone Joint Surg Am* 72(2):220–229
73. Yu S, Houghton VM, Sether LA, Ho KC, Wagner M (1989) Criteria for classifying normal and degenerated lumbar intervertebral disks. *Radiology* 170(2):523–536
74. Ziv I, Maroudas C, Robin G, Maroudas A (1993) Human facet cartilage: swelling and some physicochemical characteristics as a function of age. Part 2: age changes in some biophysical parameters of human facet joint cartilage. *Spine* 18(1):136–146