

# Powtarzalność i odtwarzalność wyników oceny rezonansu magnetycznego kręgosłupa lędźwiowego u kandydatów na pilotów samolotów

## Interrater and Intrarater Reliability of Lumbar Spine Magnetic Resonance Assessment of Military Aircraft Pilot Candidates

Aleksandra Truszczyńska<sup>1,2(C,D,E,F)</sup>, Adrian Brychcy<sup>1(D,E,F)</sup>, Kazimierz Rapała<sup>1,2(A,E,F)</sup>, Piotr Walczak<sup>1(D)</sup>, Olaf Truszczyński<sup>3(B)</sup>, Adam Tarnowski<sup>3,4(C)</sup>

<sup>1</sup> Samodzielny Publiczny Szpital Kliniczny im. prof. A. Grucy, Otwock

<sup>2</sup> Wydział Wychowania Fizycznego i Sportu w Białej-Podlaskiej, Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie

<sup>3</sup> Wojskowy Instytut Medycyny Lotniczej, Warszawa

<sup>4</sup> Wydział Psychologii, Uniwersytet Warszawski

<sup>1</sup> Professor Adam Gruca Teaching Hospital in Otwock

<sup>2</sup> Josef Pilsudski University of Physical Education in Warsaw, Faculty of Physical Education in Biała Podlaska

<sup>3</sup> Military Institute of Aviation Medicine

<sup>4</sup> Psychology Faculty, Warsaw University

### STRESZCZENIE

**Wstęp.** Rozwój diagnostyki obrazowej ułatwia ocenę anatomii kręgosłupa, natomiast nie ma pełnej zgodności pomiędzy stwierdzanymi w badaniach obrazowych, zmianami morfologicznymi i patologicznymi kręgosłupa a objawami klinicznymi.

**Cel badań.** 1. Ocena stopnia intensywności sygnału i zmian zwyrodnieniowych krążka międzykręgowego w obrazie MR kręgosłupa lędźwiowego. 2. Ocena zgodności klasyfikacji zmian przez doświadczonego i niedoświadczonego lekarza.

**Materiał i metody.** Materiał stanowiło 89 osób w wieku 18-21 lat. Obraz MR oceniano za pomocą konsoli do analizy cyfrowej obrazu przez trzech lekarzy o różnym stopniu doświadczenia zawodowego.

**Wyniki.** Zmiany w uwodnieniu krążka międzykręgowego występowały u 3,37% – 4,49% badanych na poziomie L4-L5, i u 13,48%-15,73% badanych na poziomie L5-S1. Odtwarzalność wyników wahała się od pełnej zgodności w segmentach, w których rzadko obserwowane są zmiany zwyrodnieniowe do zgodności umiarkowanej i słabej na poziomach L4-L5 i L5-S1. Powtarzalność wyniosła od pełnej zgodności na poziomach wyższych do zgodności dobrej i bardzo dobrej na poziomach L4-L5 i L5-S1. Odtwarzalność wyników dotycząca wielkości zmian w krążku międzykręgowym na poziomach Th12 do L4 wyniosła 0,6; natomiast na poziomach L4-L5 i L5-S1 zgodność była słaba. Powtarzalność wyników wahała się od bardzo dobrej zgodności na poziomie Th12 do L3, do słabej na L3-L4-L5 i średniej na poziomie L5-S1.

**Wnioski:** 1. Bezobjawowe zmiany krążka międzykręgowego stwierdzono u 30% badanych. 2. Ocena stopnia uszkodzenia krążka międzykręgowego wymaga specjalistycznego szkolenia i doświadczenia badającego.

**Słowa kluczowe:** rezonans magnetyczny, choroba dyskowa, piloci samolotów

### SUMMARY

**Background.** The development of diagnostic imaging facilitates evaluation of spinal anatomy, but there is no full correspondence between morphologic and pathologic changes observed in radiographic studies and clinical findings.

**The aim of the study was to:** 1. Evaluate the degree of signal intensity and degenerative changes of intervertebral discs in the lumbar spine on MRI. 2. Estimate interrater reliability of classification of pathology between experienced and inexperienced medical doctors.

**Material and method.** The sample consisted of 89 persons aged between 18 and 21 years. MRI scans of the lumbar spine were evaluated by the use of a console for digital analysis by three medical doctors with different levels of experience.

**Results.** Evidence of dehydration of intervertebral disc was found in 3.37%-4.49% at the L4-L5 level, and in 13.48%-15.73% of cases at the L5-S1 level. Interrater reliability ranged from full agreement at levels where degenerative changes are rarely observed to fair and poor agreement at L4-L5 and L5-S1. Intrarater reliability showed full agreement on higher levels to good and very good at L4-L5 and L5-S1 levels. Interrater reliability concerning the degree of degenerative changes was 0.6 for Th12 to L4, but was poor at L4-L5 and L5-S1. Intrarater reliability ranged from full agreement at higher levels to poor at L3-L4-L5 and moderate at L5-S1.

**Conclusions.** 1. Asymptomatic intervertebral disk changes were found in 30% of the study participants. 2. Evaluation of the degree of degenerative changes of intervertebral discs requires specialist training and experience.

**Key words:** MRI, disc disease, aircraft pilots

## WSTĘP

Prawdopodobnie u 80% osób co najmniej raz w życiu wystąpią bóle kręgosłupa lędźwiowego [1]. W USA choroba ta dotyczy 30 milionów ludzi, zaś koszty leczenia wynoszą rocznie 8 miliardów dolarów [2].

W diagnostyce obrazowej najczęściej stosowanym badaniem obrazowym jest rezonans magnetyczny. MR jest badaniem najlepiej wykazującym patologię krążka międzykręgowego. Uwidacznia on duże pole anatomiczne w płaszczyźnie czołowej i strzałkowej.

Nie ma pełnej zgodności pomiędzy stwierdzanymi w badaniach obrazowych zmianami morfologicznymi, także patologicznymi, a objawami klinicznymi. Znaczna część pacjentów ze zmianami zwyrodnieniowymi krążków międzykręgowych ma dolegliwości kliniczne: bólowe i ubytkowe, tylko w pewnym okresie, a po leczeniu np.: farmakologicznym i fizjoterapii przez wiele lat nie ma dolegliwości, pomimo tego, że nadal ma co najmniej takie same zmiany zwyrodnieniowe.

Diagnostyka MR daje jednak pewien odsetek wyników fałszywie dodatnich u osób bezobjawowych, nawet przy dokładnym badaniu klinicznym. Zmiany w obrazie MR występują u ludzi, którzy nigdy nie doświadczyli bólów kręgosłupa, co opisali m. in. Jensen i wsp. [3]. Problemem staje się również powtarzalność oceny obrazów MR.

Celem pracy była:

1. Ocena obrazów MR kręgosłupa lędźwiowego u zdrowych kandydatów na pilotów wojskowych.
2. Ocena krążka międzykręgowego uwzględniająca stopień dehydratacji, wypuklinę, przepuklinę i wypadnięcie jądra miazdzystego.
3. Zgodność odtwarzalności i powtarzalności badania przez doświadczonego i niedoświadczonego lekarza.

## MATERIAŁ I METODY

Materiał badany stanowiło 89 osób w wieku 18-21 lat. Kobiet było 21, a mężczyzn 68. Obraz MR kręgosłupa lędźwiowego zapisany na płytach CD oceniano za pomocą konsoli do analizy cyfrowej obrazu Carestream Diagnostic Workstation v. 10 VR Virtual Reading Software z opcją V3D Option (MIP, Volume Rendering, Tissue Definition). Badani informowali, że nigdy nie mieli żadnych dolegliwości ze strony kręgosłupa. Badania MR zostały przeprowadzone dla potrzeb rutynowej diagnostyki przydatności kandydatów do wysokospecjalistycznego i kosztownego szkolenia na pilotów samolotów wysokomanewrowych. Oceny obrazów MR wykonano na

## BACKGROUND

Pain in the lumbar spine occurs probably in 80% of the population at some point in their lives [1]. In the USA this ailment afflicts 30 million people and the annual cost of treatment amounts to 8 billion dollars [2].

Magnetic resonance imaging is the most popular modality of diagnostic imaging. MRI can best visualise intervertebral disc pathology. MRI scans also encompass a large anatomic area in the frontal and sagittal planes.

There is also no full correspondence between morphologic and pathologic changes observed in imaging studies and clinical findings. A significant proportion of patients with degenerative changes of intervertebral discs report pain and neurological deficits only over a certain period and after treatment (eg. pharmaco- and physiotherapy), they are free from these complaints for a number of years even though they still have degenerative changes of at least the same severity.

At the same time, magnetic resonance imaging does yield some false positive results in asymptomatic persons even with thorough clinical assessment. Changes in an MRI scan can be seen in those who have never experienced spinal pain, as described, among others, by Jensen et al. [3]. Intrarater reliability of MRI scan assessment also becomes a problem.

The aim of the study was to:

1. Evaluate MRI scans of the lumbar spine in healthy military aircraft pilot candidates.
2. Evaluate the intervertebral disc with regard to the degree of dehydration, protrusion, extrusion, or prolapse of nucleus pulposus.
3. Estimate interrater and intrarater reliability of assessment by an experienced and inexperienced doctor.

## MATERIAL AND METHODS

The study enrolled 89 participants aged between 18 and 21 years. There were 21 women and 68 men. MRI scans of the lumbar spine saved to CDs were evaluated by the use of a console for digital image analysis (Carestream Diagnostic Workstation v. 10 VR Virtual Reading Software with V3D Option [MIP, Volume Rendering, Tissue Definition]). The study participants stated that they have never had any spine-related complaints. MRI examinations were conducted as part of routine evaluation of the candidates' fitness for the highly specialised and expensive training of pilots of high-maneuvrability aircraft. The evaluation of MRI scans was requested

zlecenie wojskowego Instytutu Medycyny Lotniczej w Warszawie. Autorzy pracy przeanalizowali wykonane wcześniej w tym celu badania MR, dlatego też zgoda komisji bioetycznej nie była konieczna. Poza tym, pole magnetyczne nie jest szkodliwe, jak to jest w przypadku promieni Rtg.

Przeprowadzono 3-krotną ocenę tych samych badań na poziomach 6 poziomach: Th12-L1, L1-L2, L2-L3, L3-L4, L4-L5 i L5-S1.

Do analizy statystycznej współczynnika Kappa wykorzystano miary kierunkowe Lambda. Kappa – współczynnik Cohena mierzy zgodność ocen dwóch oceniających, gdy obaj oceniają ten sam obiekt. Wartość 1 oznacza doskonałą zgodność. Wartość 0 oznacza, że zgodność nie jest lepsza od przypadkowej. Współczynnik Kappa Cohena, określa stopień zgodności dwukrotnych pomiarów tej samej zmiennej w różnych warunkach. Pomiaru tej samej zmiennej może dokonać 2 różnych obserwatorów (odtwarzalność) lub jeden obserwator może dokonać pomiaru dwukrotnie (powtarzalność).

Sposób interpretacji:

Kappa mniej niż 0,20	– to zgodność zła
0,20- 0,40	– zgodność słaba
0,41-0,60	– zgodność średnia
0,61-0,80	– zgodność dobra
0,81- 1	– zgodność bardzo dobra.

## WYNIKI

1. W pierwszym badaniu wykonanym przez doświadczonego ortopedę (40 lat klinicznego doświadczenia ortopedycznego) oceniano krążek pod kątem: 1) stopnia dehydratacji, 2) wypukliny, 3) przepukliny, 4) wypadnięcia jądra miazdżystego. W badaniu tym uczestniczył jako osoba szkoląca się pierwszy niedoświadczony lekarz rezydent (wyniki przedstawia Tabela 1).
2. Drugie badanie samodzielnie wykonał pierwszy szkolący się lekarz rezydent po 12 miesiącach szkolenia (wyniki przedstawia Tabela 2).

by the Military Institute of Aviation Medicine in Warsaw. We analysed MRI scans performed previously for that purpose so that the approval of an ethical review board was not necessary. Unlike X-ray studies, magnetic fields are also not harmful.

MRI scans were evaluated three times at the following six levels: Th12-L1, L1-L2, L2-L3, L3-L4, L4-L5, and L5-S1.

Lambda directional measures were used for statistical analysis of the Kappa coefficient. Cohen's Kappa measures interrater reliability of evaluations made by two observers who assess the same object. A value of 1 means perfect agreement and a value of 0 means that agreement is not better than one that would be expected by chance. Cohen's Kappa coefficient defines the level of agreement between two measurements of the same variable under different conditions. The measurement of the same variable can be done either by two different observers (interrater reliability) or one observer can perform the measurement twice (intrarater reliability).

Interpretation:

A Kappa of less than 0.20 indicates	– poor agreement
0.20- 0.40	– is fair agreement
0.41-0.60	– is moderate agreement
0.61-0.80	– is good agreement
0.81- 1	– is excellent agreement.

## RESULTS

1. In the first examination, conducted by an experienced orthopaedist (40 years of clinical orthopaedic experience), discs were evaluated with regard to: 1) the level of dehydration, 2) protrusion, 3) extrusion, 4) prolapse of the nucleus pulposus. The first inexperienced resident doctor participated in this evaluation as a trainee (see Table 1 for results).
2. The second examination was conducted by the first resident doctor unassisted, after one year of training (Table 2).

Tab. 1. Zmiany w krążkach międzykręgowych w zależności od poziomu w badaniu 1 lekarza pod kierunkiem doświadczonego lekarza stwierdzone w 89 osobowej grupie badanych

Tab. 1. Changes in intervertebral discs among the 89 study participants by level in an evaluation by the first doctor under the supervision of an experienced doctor

Poziom level	Th12- L1	L1/L2	L2/L3	L3/L4	L4/L5	L5/S1
Dehydratacja dehydration	1 (1,12%)	1 (1,12%)	1 (1,12%)	2 (2,25%)	4 (4,49%)	12 (13,48%)
Wypuklina protrusion	1 (1,12%)	1 (1,12%)	0 (0%)	1 (1,12%)	8 (8,98%)	21 (23,1%)
Przepuklina extrusion	0	0 (0%)	0 (0%)	0 (0%)	1 (1,12%)	3 (3,37%)
Sekwestracja sequestration	0	0	0	0	0	0

Tab. 2. Drugie badanie przeprowadzone po rocznym przeszkoleniu ortopedy rezydenta w 89 osobowej grupie badanych

Tab. 2. A second examination conducted by the young resident orthopaedist after one year of training among the 89 study participants

Poziom Level	Th12- L1	L1/L2	L2/L3	L3/L4	L4/L5	L5/S1
Dehydratacja dehydration	0	0 (0%)	2 (2,25%)	0 (0%)	3 (3,37%)	14 (15,73%)
Wypuklina protrusion	0	0 (0%)	0 (0%)	0 (0%)	1 (1,12%)	8 (8,98%)
Przepuklina extrusion	0	0 (0%)	0 (0%)	0 (0%)	1 (1,12%)	5 (5,61%)
Sekwestracja sequestration	0	0	0	0	0	0

3. Trzecie badanie samodzielnie wykonał drugi, bardziej doświadczony (4 lata pracy) lekarz rezydent (wyniki przedstawia Tabela 3).

Oceniano zgodność wyników tego samego badającego – odtwarzalność wyników (intrarater reliability), którą porównano z oceną przeprowadzoną przez drugiego bardziej doświadczonego ortopeda – powtarzalność wyników (zgodność wyników między badającymi – interrater reliability).

Ocena uwodnienia krążka międzykręgowego (szczegółowe dane przedstawia Tabela 4).

Współczynnik wiarygodności  $\kappa$  u tego samego oceniającego dotyczący zgodności oceny tego samego badającego wahał się od pełnej zgodności w segmentach, w których rzadko obserwowane są zmiany zwyrodnieniowe (krążki jasno świecą w obrazie T2

3. The third examination was conducted by the second, more experienced (4 years of practice) resident doctor, again unassisted (Table 3).

Intrarater reliability, or the consistency of results obtained by the same observer, was evaluated and compared with the evaluation made by the other, more experienced orthopaedist for interrater reliability (between-observer consistency).

Evaluation of hydration of intervertebral disc (see Table 4 for detailed data).

The Kappa coefficient of the same observer (intrarater reliability) ranged from full agreement at levels where degenerative changes are rarely observed (discs appear bright on T2-weighted images) to fair and poor agreement at L4-L5 and L5-S1 (Table 4). As hyperintense, normal intervertebral discs at

Tab. 3. Zmiany w krążkach międzykręgowych w zależności od poziomu w badaniu przez drugiego ortopeda w 89 osobowej grupie badanych

Tab. 3. Changes in intervertebral discs among the 89 study participants by level in an evaluation by the second doctor

Poziom level	Th12- L1	L1/L2	L2/L3	L3/L4	L4/L5	L5/S1
Dehydratacja dehydration	1 (1,12%)	0 (0%)	2 (2,25%)	0 (0%)	3 (3,37%)	14 (15,73%)
Wypuklina protrusion	0	0 (0%)	1 (1,12%)	3 (3,3%)	12 (13,2%)	18 (19,8%)
Przepuklina extrusion	0	0 (0%)	0 (0%)	0 (0%)	1 (1,12%)	6 (6,6%)
Sekwestracja sequestration	0	0	0	0	0	0

Tab. 4. Diagnozowanie zdrowego (świecącego) i z cechami dehydratacji krążka międzykręgowego w 89 osobowej grupie badanych

Tab. 4. Identification of healthy (glistening) and dehydrated intervertebral discs among the 89 study participants

	Intrarater reliability	Interrater reliability
Th12-L1	Pełna zgodność – Kappa 1 Full	Pełna zgodność – Kappa 1 Full
L1-L2	Pełna zgodność – Kappa 1 Full	Pełna zgodność – Kappa 1 Full
L2-L3	Zgodność zła – Kappa 0.14 Poor	Pełna zgodność – Kappa 1 Full
L3-L4	Pełna zgodność – Kappa 1 Full	Pełna zgodność – Kappa 1 Full
L4-L5	Zgodność zła – Kappa 0.14 Poor	Zgodność dobra – Kappa 0.7 Good
L5-S1	Zgodność umiarkowana – Kappa 0.28 Moderate	Zgodność bardzo dobra – Kappa 0.94 Very good

Tab. 5. Zgodność w diagnozowaniu zmian w postaci wypukliny, przepukliny jądra miazdżystego w 89 osobowej grupie badanych  
 Tab. 5. Intra- and interrater reliability in diagnosis of protrusion or extrusion of the nucleus pulposus among the 89 study participants

	Intrarater reliability	Interrater reliability
Th12-L1	Zgodność średnia/dobra – Kappa 0.6 Average/good	Zgodność bardzo dobra – Kappa 1 Excellent
L1-L2	Zgodność średnia/dobra – Kappa 0.6 Average/good	Zgodność bardzo dobra – Kappa 1 Excellent
L2-L3	Zgodność średnia/dobra – Kappa 0.6 Average/good	Zgodność bardzo dobra – Kappa 1 Excellent
L3-L4	Zgodność średnia/dobra – Kappa 0.6 Average/good	Zgodność słaba – Kappa 0.28 Excellent
L4-L5	Zgodność słaba – Kappa 0.26 Poor	Zgodność słaba – Kappa 0.21 Excellent
L5-S1	Zgodność słaba – Kappa 0.26 Poor	Zgodność średnia – Kappa 0.46 Excellent

zależnym) do zgodności umiarkowanej i słabej na poziomach L4-L5 i L5-S1 (Tab. 4). Hiperintensywne, prawidłowe krążki międzykręgowe na wyższych poziomach były łatwe do oceny, dlatego Kappa wykazał pełną zgodność.

Na dwóch najniższych poziomach obserwowano zmiany w uwodnieniu krążka międzykręgowego. Występowały one u 3,37 – 4,49% badanych na poziomie L4-L5, i u 13,48 – 15,73% badanych na poziomie L5-S1.

Natomiast porównanie wyników pomiędzy badającymi wahało się od pełnej zgodności na poziomach wyższych do zgodności dobrej i bardzo dobrej na poziomach najniższych (L4-L5 i L5-S1).

Ocena wielkości zmian krążka międzykręgowego (szczegółowe dane przedstawia tabela 5)

Odtwarzalność wyników na poziomach Th 12 do L4 wyniosła 0,6 (na granicy średniej i dobrej zgodności), natomiast na poziomach L4-L5 i L5-S1 zgodność była słaba (0,26).

Powtarzalność wyników wahała się już od bardzo dobrej zgodności na poziomie Th12 do L3, do słabej na L3-L4-L5 i średniej Kappa 0,46 na poziomie L5-S1.

## DYSKUSJA

Uzyskane wyniki świadczyły, że brak zmian w uwodnieniu krążka międzykręgowego jest łatwy w ocenie zarówno przez niedoświadczonego, jak i przez doświadczonego lekarza. Współczynnik Kappa wykazał pełną zgodność. Natomiast brak doświadczenia utrudniał ocenę stopnia uwodnienia krążka, dlatego też zgodność była niska lub umiarkowana. Zdobycie doświadczenia spowodowało znaczną poprawę zgodności.

Poprawiła się również znacznie zgodność oceny pomiędzy badającymi dotycząca stopnia wypukliny i przepukliny. Najniższa – słaba – zgodność pomiędzy badającym występowała na dwóch najniższych poziomach.

higher levels were easy to evaluate, there was full agreement by Kappa.

Evidence of dehydration of intervertebral disc was found at the two lowest levels. These changes were found in 3.37–4.49% of the study participants at L4-L5, and in 13.48%–15.73% at L5-S1.

The comparison of the results between the observers ranged from full agreement at the higher levels to good and very good agreement at the lowest levels (L4-L5 and L5-S1).

Evaluation of the degree of degenerative changes (see Table 5 for detailed data).

Interrater reliability at the levels from Th12 to L4 was 0.6 (borderline between moderate and good agreement), but at the levels of L4-L5 and L5-S1 the agreement was fair (0.26).

Intrarater reliability ranged from full agreement at the levels from Th12 to L3, to poor at L3-L4-L5 and a moderate Kappa of 0.46 at L5-S1.

## DISCUSSION

Our results show that the absence of changes in the hydration of intervertebral discs is easily evaluated by both an experienced and inexperienced doctor. The Kappa coefficient showed full agreement. However, lack of experience caused difficulties in the evaluation of the degree of disc hydration, which resulted in fair or moderate agreement. Gaining experience brought about a significant improvement in agreement.

Between-observer agreement on the degree of protrusion and extrusion also improved significantly. The lowest (poor) agreement between the observers was found for the two lowest levels.

W polskiej literaturze ukazała się jedna praca opisująca bezobjawowe zmiany w kręgosłupie w obrazie MR u 98 kandydatów na pilotów wojskowych opublikowana przez Rapala A i wsp. Uzasadnieniem wykonania badań rezonansu był wysoki koszt szkolenia pilotów samolotów odrzutowych. Dehydratację krążka międzykręgowego stwierdzono u 23,5% badanych, a 22% spośród tej grupy miało wypuklinę krążka międzykręgowego [4]. Takich badań u kandydatów na pilotów w literaturze anglojęzycznej autorzy pracy nie znaleźli.

W badaniach MR przeprowadzonych przez Weinreb i wsp. porównano występowanie zmian w MR kręgosłupa lędźwiowego u 45 kobiet w ciąży i 41 kobiet w wieku rozrodczym, ale nie będącymi w ciąży. Zmiany pod postacią wypukliny i przepukliny występowały u 53% ciężarnych i 54% z grupy kontrolnej. Autorzy sugerowali, że zmiany w krążkach międzykręgowych są typowe dla kobiet w tym przedziale wieku [5].

Wiesel i wsp. badając za pomocą TK 52 osoby z bólami kręgosłupa piersiowo-lędźwiowego stwierdzili, że przepuklina na poziomie L4/L5/S1 występowała u 19,5% u osób poniżej 40 lat i 26,9% u osób powyżej 40 lat [6].

Również Szczerbo-Trojanowska za Marchiorim podaje, że bezobjawowe zmiany w obrazie MR dotyczą pomiędzy 20-30% populacji [7].

Z badań Bodena i wsp. wynika, że zmiany w obrębie kręgosłupa w obrazie MR są ściśle związane z wiekiem. U 67 osób z bólami krzyża stwierdzono przepuklinę w 20% u ludzi poniżej 60 r.ż. i u 36% u osób powyżej 60 r.ż. U 20% badanych wieku 20-29 lat stwierdzono wypuklinę u 20-35%, a przepuklinę od 15-25% [8].

Wiesel i wsp. oraz Boden i wsp. sugerowali nawet, że istnieje ryzyko wykonania operacji na podstawie niepokojących zmian w obrazie MR, nie będących przyczyną bólów [6,8]. W takich przypadkach decydujące jest dokładne badanie kliniczne. Zmiany w krążku międzykręgowym narastają z upływem lat i jest to naturalny proces starzenia, co potwierdziły badania Modica [9].

Borenstein i wsp. przeprowadzili badania polegające na ocenie obrazów MR u 67 osób, którzy nie mieli żadnych dolegliwości ze strony kręgosłupa piersiowo-lędźwiowego. W 31% badanych obrazów stwierdzono nieprawidłowy krążek międzykręgowy lub kanał kręgowy. Po upływie 7 lat 58% badanych wciąż nie miało żadnych dolegliwości ze strony kręgosłupa piersiowo-lędźwiowego, pomimo zwiększenia stopnia zmian zwyrodnieniowych [10].

Badania w dużej grupie młodych ludzi przeprowadzili Takatalo i wsp. Przeanalizowali badania MR u 874 osób w wieku 18-21 lat. Stwierdzili oni, że bó-

Only one study by Polish authors (Rapala A et al.) has been published that describes asymptomatic changes in the spine on MRI scans of 98 military aircraft pilot candidates. MRI examinations were performed because of the high costs of training of jet aircraft pilots. Dehydration of intervertebral disc was found in 23.5% of the study participants, and 22% had intervertebral disc protrusion [4]. We did not find similar reports of examinations in pilot candidates in the English-language literature.

The prevalence of abnormalities in lumbar spine MRIs of 45 pregnant women and 41 non-pregnant women of childbearing age was compared in an MRI-based study conducted by Weinreb et al. Abnormalities in the form of disc protrusion and extrusion were found in 53% of the pregnant women and 54% of the controls. The authors suggested that intervertebral disc pathology is typical of women at this age [5].

Wiesel et al. carried out a CT evaluation of 52 persons with thoracolumbar spine pain, finding disc extrusion at L4/L5/S1 in 19.5% of those participants under the age of 40 and 26.9% of the patients over 40 years of age [6].

Citing Marchiori, Szczerbo-Trojanowska also states that asymptomatic changes on MRI scans are seen in approximately 20-30% of the population [7].

A study by Boden et al. indicates that spinal pathology visible on MRI scans is closely related to age. In a group of 67 persons with low back pain, disc extrusion was found in 20% of those under the age of 60 and in 36% of those over 60 years of age. Among the 20% aged 20-29, disc protrusion was found in 20-35% and extrusion in 15-25% [8].

Wiesel et al. as well as Boden et al. actually suggested that there is a danger of performing surgery on the basis of finding alarming MRI abnormalities that are not the cause of pain [6,8]. A thorough clinical examination is crucial in such cases. Intervertebral disc pathology becomes more marked over time and, as confirmed by Modic, this is a natural ageing process [9].

Borenstein et al. evaluated MRI scans of 67 persons with no thoracolumbar spine-related complaints. Intervertebral disc or vertebral canal abnormalities were found on 31% of the scans. After 7 years 58% of the group still did not have any thoracolumbar spine symptoms, even though the degree of degenerative changes had increased [10].

Takatalo et al. studied a large group of young individuals. They analysed MRI scans of 874 persons between the ages of 18 and 21 and found that spine pain correlated with MRI abnormalities, but changes in intervertebral discs were also found in 30% of asymptomatic participants [11].

le kręgosłupa korelowały ze zmianami w MR, natomiast u 30% bezobjawowych osób również występowały zmiany w krążkach [11].

W innej pracy 50% badanej młodzieży miało zmiany zwyrodnieniowe w obrębie co najmniej jednego krążka, a 25% miało wypuklenie jądra miazdżystego [12].

Chińscy autorzy przeanalizowali obraz MR u 1043 osób w wieku 18-55 lat. Zmiany w krążkach międzykręgowych występowały u 45% osób poniżej 30 roku życia i wzrastały wraz z wiekiem [13].

Porównywano również wyniki dyskograficzne zmian w krążku międzykręgowym u bezobjawowych ochotników z grupą pacjentów z przewlekłym bólem kręgosłupa. Autorzy stwierdzili, że pacjenci z objawami choroby dyskowej z przerwaniem pierścienia włóknistego wykazują niższą tolerancję bólu podczas dyskografii w porównaniu z osobami grupy kontrolnej, którzy nie mają dolegliwości, a mają przerwany pierścień włóknisty [14].

Interesujące badania przeprowadzili Brant-Zawadzki i wsp. Porównywali oni dwie terminologie opisujące nieprawidłowości w obrazie MR kręgosłupa lędźwiowego. W grupie I oceniano krążek międzykręgowy pod względem budowy – norma, wypuklina, przepuklina, w grupie II ocena dotyczyła normy, wypukliny, przepukliny i wypadnięcia. Zgodność pomiędzy badaczami  $\kappa$  wynosiła 0,58, powtarzalność u jednego badacza 0,86, współczynnik  $\kappa = 0,71$  u pierwszego badacza, a u drugiego  $\kappa = 0,69$ . Doświadczenia z odczytaniem standardowej terminologii wykazały umiarkowaną zgodność w interpretacji przemieszczania się dysku w przestrzeni międzykręgowej w obrazie MR [15].

W publikacji autorów Chiodo i wsp. porównywano badanie MR i EMG u 35 osób z bezobjawowymi zmianami MR. Po 18 miesiącach powtórzono badanie, z którego wynikało, że badanie EMG ma mniejszą liczbę fałszywie dodatnich wyników, niż MR u bezobjawowych dorosłych w ocenie stenozy lędźwiowej [16].

Bardzo ciekawe badania przeprowadzili Savage i wsp. Autorzy oceniali zależność pomiędzy występowaniem bólów kręgosłupa, a wiekiem, rodzajem wykonywanej pracy oraz obrazem MR. Zbadano 149 mężczyzn w wieku 20-58 lat. Okazało się, że nie było zależności istotnej statystycznie pomiędzy obrazem MR, a występowaniem bólów kręgosłupa. W grupie bezobjawowych osób 32% miało zmiany w MR, natomiast w grupie osób z bólami kręgosłupa 47% miało prawidłowy obraz MR [17].

Z badań Mileckiego i wsp. zgodność obrazu MR ze zmianami stwierdzanymi śródoperacyjnie u 180 chorych wynosiła 96% [18].

Znaczenie współpracy pomiędzy radiologami i chirurgami podkreślili w swoim artykule Brand-

In another study, degenerative changes in at least one disc were found in 50% of the young people examined and 25% of them had a protrusion of the nucleus pulposus [12].

Chinese authors analysed MRI scans of 1043 persons aged 18-55 years. Changes in intervertebral discs were found in 45% of the part of the sample under 30 years of age, and they progressed with advancing age [13].

In another study, discography data showing intervertebral disc changes in asymptomatic volunteers were compared with those of a group of patients with chronic spinal pain. The authors found that patients with symptoms of disc disease and annular disruption demonstrated lower pain tolerance during discography compared to controls who had annular disruption but no pain [14].

Brant-Zawadzki et al. conducted an interesting study in which they compared two terminologies describing abnormalities on lumbar spine MRI scans. In group 1, intervertebral discs were classified structurally as 'normal disc', 'protrusion' and 'extrusion'; in group 2 the assessment specified the following categories: normal disc, protrusion, extrusion, and prolapse. Between-observer consistency  $\kappa$  was 0.58, intrarater reliability in one observer was 0.86, Kappa coefficient ( $\kappa$ ) of the first observer was 0.71, and Kappa coefficient ( $\kappa$ ) of the second observer was 0.69. This study of understanding of standard terminology showed moderate agreement on interpretation of intervertebral disc displacement within the intervertebral space on MRI scans [15].

Chiodo et al. compared MRI and EMG results of 35 persons with asymptomatic changes seen on MRI scans. The evaluation was repeated after 18 months and it turned out that EMG produced a smaller number of false positive results than MRI when used for assessment of lumbar spine stenosis in asymptomatic adults [16].

Savage et al. conducted a very interesting study evaluating the correlation between the incidence of spinal pain and age, type of work performed, and MRI findings. They examined 149 men between the ages of 20 and 58 years. In that study, there was no statistically significant correlation between MRI findings and the incidence of spinal pain. MRI abnormalities were detected in 32% of asymptomatic participants, while 47% of those with spinal pain had normal MRI scans [17].

According to a study conducted by Milecki et al. in a group of 180 patients, the degree of consistency of MRI findings with abnormalities diagnosed intraoperatively amounted to 96% [18].

Zawadzki i wsp. Często przyczyny bólów kręgosłupa, pomimo nieprawidłowości w obrazie MR, nie znajdują potwierdzenia w dokładnych badaniach klinicznych. Podkreślają również, że nie ma wystarczającej standaryzacji terminologii radiologicznej, jak i postępowania chirurgicznego [19].

Ważny problem został poruszony przez autorów Ash i wsp. Analizowali oni wpływ informacji, którą uzyskiwali pacjenci dotyczące ich wyników badania MR. Jako hipotezę badawczą postawili tezę, że informacja uspokoi chorych. Przebadali oni 246 pacjentów z ostrymi bólami kręgosłupa lędźwiowego i/lub radikulopatią leczonych zachowawczo. W grupie badanej wykonano MR, zaś w kontrolnej nie wykonano badania MR. Wyniki wykazały, że nie było różnicy istotnej statystycznie w wynikach leczenia, natomiast grupa badana miała gorsze samopoczucie [20].

Zmiany w krążkach międzykręgowych oceniać można również w TK. Krupski i wsp. stwierdzili u 60 chorych z bólami przewlekłymi występowanie zmian w krążkach międzykręgowych na poziomie L5-S1 w 50 przypadkach, L4-L5 w 45 przypadkach i na poziomie L3-L4 w 15 [21].

## WNIOSKI

1. Obniżenie sygnału i zmiany zwyrodnieniowe krążków międzykręgowych w obrazie MR u młodych, zdrowych ludzi stwierdzono u 30% badanych.
2. Ocena zdrowych krążków międzykręgowych wykazała pełną zgodność zarówno w badaniu powtarzalnym (intra-rater reliability), jak w odtwarzalnym (inter-rater reliability).
3. W ocenie stopnia uszkodzenia krążka międzykręgowego niezbędne jest specjalistyczne szkolenie i doświadczenie badającego.

## PIŚMIENNICTWO / REFERENCES

1. Robertson JT. The rape of the spine. *Surgical Neurology*. 1993; 39: 5-12.
2. Deyo RA, Tsuiwu YJ. Descriptive epidemiology of low-back-pain and its related medical-care in the united-states. *Spine*. 1987; 12: 264-268.
3. Jensen MC, Brant- Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic-resonance-imaging of the lumbar spine in people without back pain. *New England Journal of Medicine*. 1994; 331: 69-73.
4. Rapala A, Rapala K, Łukawski St. Nonsymptomatic changes in magnetic resonance imaging of spine among airforce candidates. *Chir Narzadow Ruchu Ortop Pol*. 2005; 70: 411-414.
5. Weinreb JC, Wolbarsht LB, Cohen JM, Brown CEL, Maravilla KR. Prevalence of lumbosacral intervertebral-disk abnormalities on MR images in pregnant and asymptomatic nonpregnant women. *Radiology*. 1989; 170: 125-128.
6. Wiesel SW, Tsourmas N, Feffer HL, Citrin CM, Patronas N. 1984 volvo award in clinical sciences - a study of computer-assisted tomography .1. The incidence of positive cat scans in an asymptomatic group of patients. *Spine*. 1984; 9: 549-551.
7. Marchiori D. M.: *Radiologia kliniczna*. Wyd. Czelej, Kraków, 1999: 532.
8. Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects - a prospective investigation. *Journal of Bone and Joint Surgery-American Volume*. 1990; 72A: 403-408.
9. Modic MT. Degenerative disc disease and back pain. *Magn Reson Imaging Clin N Am*. 1999; 7: 481-491.
10. Borenstein DG, O'Mara JW, Boden SD i wsp. The value of magnetic resonance imaging of the lumbar spine to predict low-back pain in asymptomatic subjects - a seven-year follow-up study. *Journal of Bone and Joint Surgery-American Volume*. 2001; 83A: 1306-1311.

Brand-Zawadzki et al. emphasize the importance of collaboration between radiologists and surgeons. Despite abnormal findings on MRI scans, careful clinical examinations do not confirm frequent causes of spinal pain. The researchers also underline the lack of adequate standardization of both radiographic terminology and surgical management plans [19].

Ash et al. discuss an important issue. They analysed how information concerning MRI results influenced patients. Their research hypothesis was that the information would make patients calmer. The researchers examined 246 patients with acute lumbar spine pain and/or radiculopathy undergoing conservative treatment. MRI was performed in the experimental group but not in the control group. The results demonstrated no statistically significant difference in treatment results, but the experimental group demonstrated poorer overall well-being [20].

Changes in intervertebral discs can be also evaluated on CT scans. Krupski et al. examined 60 patients with chronic pain and found evidence of intervertebral disc changes at L5-S1 in 50 cases, at L4-L5 in 45 cases, and at L3-L4 in 15 cases [21].

## CONCLUSIONS

1. Decreased signal intensity and degenerative lesions of intervertebral discs on MRI scans of young, healthy persons were found in 30% of the study participants.
2. Evaluation of healthy intervertebral discs demonstrated full agreement with regard to both intra-rater and inter-rater reliability.
3. Evaluation of the degree of degenerative changes of intervertebral discs requires specialist training and experience.



11. Derby R, Kim BJ, Lee SH, Chen Y, Seo KS, Aprill C. Comparison of discographic findings in asymptomatic subject discs and the negative discs of chronic lbp patients: Can discography distinguish asymptomatic discs among morphologically abnormal discs? *Spine j. United States* 2005; 389-394.
12. Takatalo J, Karppinen J, Niinimäki J i wsp.. Does lumbar disc degeneration on MRI associate with low back symptom severity in young Finnish adults? *Spine* 2011 Feb 24. [Epub ahead of print].
13. Takatalo J, Karppinen J, Niinimäki J, Taimela S, Näyhä S, Järvelin MR, Kyllönen E, Tervonen O. Prevalence of degenerative imaging findings in lumbar magnetic resonance imaging among young adults. *Spine* 2009;15;34:1716-21.
14. Cheung KM, Karppinen J, Chan D i wsp. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. *Spine* 2009; 20;34(9): 934-940.
15. Brant-Zawadzki MN, Jensen MC, Obuchowski N, Ross JS, Modic MT. Interobserver and intraobserver variability in interpretation of lumbar disc abnormalities - a comparison of 2 nomenclatures. *Spine*. 1995; 20: 1257-1263.
16. Chiodo A, Haig AJ, Yamakawa KSJ, Quint D, Tong H, Choksi VR. Needle emg has a lower false positive rate than mri in asymptomatic older adults being evaluated for lumbar spinal stenosis. *Clinical Neurophysiology*. 2007; 118: 751-756.
17. Savage RA, Whitehouse GH, Roberts N. The relationship between the magnetic resonance imaging appearance of the lumbar spine and low back pain, age and occupation in males. *Eur Spine J*. 1997;6:106-14.
18. Milecki M, Łukawski St, Białecki J, Sosnowski R, Lachowicz W.: Wartość metod diagnostycznych w rozpoznawaniu choroby dyskowej kręgosłupa lędźwiowego *Ortop, traumat, rehab* 2004; 6:155-159.
19. Brant-Zawadzki MN, Dennis SC, Gade GF, Weinstein MP. Low back pain. *Radiology*. 2000; 217:321-330.
20. Ash LM, Modic MT, Obuchowski NA, Ross JS, Brant-Zawadzki MN, Grooff PN. Effects of diagnostic information, per se, on patient outcomes in acute radiculopathy and low back pain. *AJNR Am J Neuroradiol*. 2008 Jun;29(6):1098-103.
21. Krupski W, Majcher P, Tataro M, Fatyga M. Diagnostyka TK schorzeń krążka międzykręgowego i zmian zwyrodnieniowo-wytwórczych kręgosłupa lędźwiowo-krzyżowego. *Ortop, traumat, rehab* 2004; 6(2):160-165.

**Liczba słów/Word count:** 5107

**Tabele/Tables:** 5

**Ryciny/Figures:** 0

**Piśmiennictwo/References:** 21

*Adres do korespondencji / Address for correspondence*

*prof. dr hab. Kazimierz Rapala*

*01-839 Warszawa, ul. Barcicka 47*

*tel. 667-060-066, e-mail. [aleksandra.rapala@wp.pl](mailto:aleksandra.rapala@wp.pl)*

*Otrzymano / Received*

*28.05.2011 r.*

*Zaakceptowano / Accepted*

*29.07.2011 r.*



*This copy is for personal use only - distribution prohibited.*

