

Szymon Pietrzak <sup>1</sup>, Marek Napiontek <sup>2</sup>, Marek Tomaszewski <sup>2</sup><sup>1</sup> Klinika Ortopedii CMKP, Otwock<sup>2</sup> Katedra i Klinika Ortopedii i Traumatologii Dziecięcej Akademii Medycznej im. K. Marcinkowskiego, Poznań

# Klasyfikacje radiologiczne Catteralla i Herringa w ocenie choroby Perthesa – zgodność i powtarzalność oceny

## *Catterall and Herring classifications in assessing Perthes disease – inter- and intra-observer study*

**Słowa kluczowe:** choroba Perthesa, klasyfikacje Catteralla i Herringa, staw biodrowy  
**Key words:** Perthes disease, Catterall and Herring classification, hip joint

### SUMMARY

**Background.** The purpose of this study was to compare the inter-observer validity and intra-observer reliability of the Catterall and Herring classification systems used in assessing pathological changes in Perthes' disease.

**Material and methods.** The material consisted in x-rays from 58 patients (12 girls and 46 boys, 63 hips, average age 6) treated conservatively or surgically. The x-rays were assessed by a board-certified orthopedist with 19 years of experience, an orthopedist just taking board exams with 5 years of experience, and a resident. For statistical analysis we used the weighted kappa coefficient and the percentage of agreement.

**Results.** The majority of the cases were qualified to groups II and III, and the younger observers evaluated more severely. The least inter-observer validity in both classification systems occurred between the resident and the senior physician. When the material was divided into two age groups (younger and older than 6), the validity of the Catterall scheme was similar in both groups, while in the Herring classification there was greater inter-observer validity for the older children. The Herring system shows greater intra-observer reliability overall; in the two age groups, the Catterall system shows greater reliability in the younger children, and the Herring system in the older.

**Conclusions.** The Catterall classification system shows less validity and reliability than the Herring system, especially in reference to children older than 6. The greatest difficulties in classification decisions occur between Catterall groups II and III.

### STRESZCZENIE

**Wstęp.** Celem pracy jest porównanie poziomu zgodności i powtarzalności kwalifikacji zmian chorobowych według klasyfikacji Catteralla i Herringa w trakcie leczenia choroby Perthesa.

**Materiał i metody.** Materiał obejmował radiogramy 58 chorych (12 dziewcząt i 46 chłopców, 63 stawy biodrowe) leczonych nieoperacyjnie lub zakwalifikowanych do leczenia operacyjnego. Średni wiek chorych wynosił 6 lat. W skład zespołu oceniającego radiogramy wchodził: specjalista II stopnia z 19-letnim stażem pracy, specjalizujący się ortopeda z 5-letnim stażem pracy oraz lekarz odbywający staż podyplomowy. W analizie statystycznej zastosowano ważony współczynnik kappa oraz odsetek zgodności ocen.

**Wyniki.** Największą liczbę przypadków zakwalifikowano do grupy II i III, a obserwatorzy młodszy kwalifikowali ostrzej. Najmniejsza zgodność ocen według każdej z klasyfikacji występuje pomiędzy lekarzem stażystą i najstarszym z obserwatorów. Po podziale materiału na dwie grupy wiekowe (poniżej i powyżej 6 roku życia) dla klasyfikacji Catteralla zgodność ocen w obu grupach była podobna. Dla klasyfikacji Herringa – zgodność ocen była większa dla dzieci starszych. Stwierdzono ponadto, że system Herringa cechuje się większą pow-

tarzalnością ocen niż system klasyfikacji Catteralla. Po podziale materiału na dwie grupy wiekowe okazuje się, że powtarzalność oceny w klasyfikacji Catteralla jest większa w grupie dzieci młodszych, natomiast klasyfikacji Herringa – dla dzieci starszych.

**Wnioski.** Klasyfikacja Catteralla cechuje się mniejszą powtarzalnością i zgodnością oceny niż klasyfikacja Herringa, zwłaszcza w odniesieniu do dzieci powyżej 6 roku życia oraz że największe trudności ocenianym sprawia rozróżnienie grupy II i III w klasyfikacji Catteralla.

## WSTĘP

Choroba Perthesa (łac. osteochondrosis coxae juvenilis, morbus Legg – Calvé – Waldenström – Perthes) – jałowa martwica głowy kości udowej u dzieci, jest jednym z bardziej znanych schorzeń narządu ruchu. Poświęcono jej wiele publikacji dotyczących etiologii, patogenez, obrazowania i metod leczenia. Znajomość historii naturalnej oraz wieloletnie doświadczenia w leczeniu tego schorzenia doprowadziły do opracowania wielu systemów oceny radiologicznej, na podstawie których kwalifikuje się chorego do jednego ze sposobów leczenia nieoperacyjnego lub operacyjnego. Klasyfikacje zmian radiologicznych opracowane przez A. Catteralla [1] oraz A. Herringa [2] stosowane są powszechnie w celu oceny rozległości zajęcia głowy kości udowej i doboru odpowiedniej metody leczenia oraz w prognozowaniu jego wyniku. Jednakże żaden z opracowanych systemów szeregujących zmiany w stawie biodrowym nie został uznany za doskonały.

Celem pracy jest porównanie wartości ww. klasyfikacji na podstawie zgodności i powtarzalności oceny dokonywanej według kryteriów, na których obie te klasyfikacje są zbudowane.

## MATERIAŁ I METODY

Materiał obejmował radiogramy 58 chorych (63 stawy biodrowe) leczonych nieoperacyjnie lub zakwalifikowanych do leczenia operacyjnego w Klinice Ortopedii Dziecięcej AM w Poznaniu w latach 1990-1994. Ocenie poddano radiogramy wykonane w pierwszych 6 miesiącach choroby. Wiek chorych wahał się od 1 roku i 9 miesięcy do 12 lat i 9 miesięcy (średnia 6 lat). W badanej grupie było 12 dziewcząt i 46 chłopców. U 24 chorych schorzenie dotyczyło lewego, u 29 chorych – prawego, a u 10 chorych obu stawów biodrowych.

W odniesieniu do klasyfikacji Catteralla (bez uwzględniania czynników ryzyka) i klasyfikacji Herringa przeprowadziliśmy badanie dotyczące zgodności oceny (ang. inter-observer study) i powtarzalności oceny (intra-observer study). Oceniano radiogramy w projekcji przednio – tylnej w rotacji 0 oraz w projekcji osiowej. W skład zespołu ocenianego wchodził:

- specjalista II stopnia z 19-letnim stażem pracy (obserwator 1)
- specjalizujący się ortopeda z 5-letnim stażem pracy (obserwator 2)
- lekarz odbywający staż podyplomowy (obserwator 3).

W części dotyczącej zgodności oceny wszyscy trzej oceniali te same radiogramy niezależnie od siebie. W części dotyczącej powtarzalności oceny te same radiogramy oceniane były przez każdego z obserwatorów w odstępie 12 miesięcy. Ponadto dla oceny statystycznej materiał podzielono na dwie grupy wiekowe – poniżej i powyżej 6 roku życia.

W analizie statystycznej zastosowaliśmy ważony współczynnik kappa oraz odsetek zgodności ocen [3]. Współczynnik ten jest wyrażony w skali od 0 do 1. Wartości poniżej 0,4 oznaczają słabą zgodność; od 0,4 do 0,75 – zgodność zadowalającą i dobrą; wartości powyżej 0,75 – zgodność wysoką.

## WYNIKI

Liczbę przypadków przydzielonych do każdej z grup według klasyfikacji Catteralla i klasyfikacji Herringa przedstawia Tabela 1.

Wynika z niej, że największą liczbę przypadków zakwalifikowano do grupy II i III wg Catteralla, a obserwatorzy młodszy kwalifikowali ostrzej, tzn. z przewagą grupy III i IV nad pozostałymi.

Zgodność pomiędzy poszczególnymi obserwatorami dla tej klasyfikacji przedstawia Tabela 2.

Wynika z niej, że największa zgodność ocen występuje pomiędzy obserwatorem 2 i 3 oraz 1 i 2, a najmniejsza pomiędzy 1 i 3. Zgodność oceny pozostaje odpowiednio na poziomie zadowalającym (obserwator 1 i 2 oraz 2 i 3) i słabym (obserwator 1 i 3). Dla grupy II i III według klasyfikacji Catteralla liczba ocen zgodnych jest najbardziej zbliżona do tzw. rozkładu oczekiwanego, tzn. sytuacji, gdyby oceny przydzielane były przypadkowo. Pomimo że w tych przypadkach byliśmy zgodni, nie ma to znaczenia statystycznego.

Tabela 3 przedstawia analogicznie wyniki zgodności ocen dla klasyfikacji Herringa. W tym przypadku zgodność ocen dla każdej z par obserwatorów jest dobra.

Po podziale materiału na dwie grupy wiekowe (poniżej i powyżej 6 roku życia) dla klasyfikacji Catteralla zgodność ocen w obu grupach była podobna. W klasyfikacji Herringa – zgodność ocen była większa dla dzieci starszych (Tab. 4).

W podobny sposób przeprowadziliśmy badania powtarzalności oceny wobec klasyfikacji Catteralla i Herringa (Tab. 5 i 6). Stwierdziliśmy, że system stworzony przez Herringa cechuje się większą powtarzalnością ocen, niż system klasyfikacji Catteralla.

Po podziale materiału na dwie grupy wiekowe okazało się, że powtarzalność oceny w klasyfikacji Catteralla jest większa w grupie dzieci młodszych, natomiast klasyfikacji Herringa – dla dzieci starszych. Z szablonu tego wyłamały się jedynie wyniki oceny najmłodszego z obserwatorów (obserwator 3) – o najmniejszym doświadczeniu zawodowym, którego powtarzalność oceny nie osiąga zakresu zadowalającego w przypadku radiogramów dzieci powyżej 6 roku życia (Tab. 7).

Tab. 1. Rozkład ocen według klasyfikacji Catteralla i klasyfikacji Herringa

Tab. 1. The distribution of assessment according to Herring and Catterall classifications

Obserwator	Grupa wg Catteralla				Grupa wg Herringa		
	I	II	III	IV	A	B	C
1	8	26	22	7	14	35	14
2	5	17	24	17	11	36	16
3	7	21	24	10	8	39	15

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

Tab. 2. Zgodność oceny między poszczególnymi parami obserwatorów dla klasyfikacji Catteralla

Tab. 2. The agreement between the pairs of observers for the Catterall classification

Para obserwatorów	Zgodność każdej z grup wg Catteralla				% zgodności	Współczynnik kapp
	I	II	III	IV		
1 i 2	4 (1)*	10 (7)	9 (8)	6 (2)	46	0,42
2 i 3	4 (1)	10 (5)	12 (9)	7 (3)	53	0,49
1 i 3	5 (1)	10 (9)	7 (9)	2 (1)	39	0,28

\* liczby w nawiasach oznaczają tzw. oczekiwane liczby zgodnych ocen, gdyby rozkład ocen był przypadkowy

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

Tab. 3. Zgodność oceny między poszczególnymi parami obserwatorów dla klasyfikacji Herringa

Tab. 3. The agreement between the pairs of observers for the Herring classification

Para obserwatorów	Zgodność każdej z grup wg Herringa			% zgodności	Współczynnik kapp
	A	B	C		
1 i 2	10 (2)*	29 (20)	11 (4)	79	0,70
2 i 3	6 (1)	30 (23)	11 (4)	76	0,61
1 i 3	8 (2)	30 (22)	10 (3)	77	0,65

\* liczby w nawiasach oznaczają tzw. oczekiwane liczby zgodnych ocen, gdyby rozkład ocen był przypadkowy

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

## DYSKUSJA

Powszechnie przyjęty system oceny oparty na klasyfikacji jakościowej, który ma na celu ułatwienie wyboru metody leczenia oraz ma zawierać przesłanki odnośnie czasu trwania choroby i spodziewanych jej następstw, powinien być z założenia łatwy w zastosowaniu oraz w miarę „obiektywny”, tzn. cecho-

wać się zadowalającą powtarzalnością oceny w czasie i zgodnością oceny, jeśli przeprowadzana jest przez różne osoby. Oba systemy oceny – klasyfikacja Catteralla i klasyfikacja Herringa, stosowane są najpowszechniej podczas leczenia dzieci z chorobą Perthesa.

Klasyfikacja Catteralla, bardziej rozbudowana i uzupełniona jeszcze o czynniki ryzyka, przez wielu

Tab. 4. Zgodność oceny między poszczególnymi parami obserwatorów dla różnych grup wiekowych  
Tab. 4. The agreement between the pairs of observers for different age groups

Para obserwatorów	Klasyfikacja Catteralla				Klasyfikacja Herringa			
	% zgodności		współczynnik kappa		% zgodności		współczynnik kappa	
	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.
1 i 2	38	53	0,38	0,44	72	85	0,65	0,74
2 i 3	52	55	0,44	0,46	62	88	0,49	0,72
1 i 3	38	38	0,34	0,18	76	79	0,69	0,57

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

Tab. 5. Powtarzalność oceny dla klasyfikacji Catteralla  
Tab. 5. The intra-observer agreement for Catterall classification

Powtórzona ocena obserwatora	Zgodność każdej z grup wg Catteralla				% zgodności	Współczynnik kappa
	I	II	III	IV		
1	5 (1)*	17 (11)	13 (8)	2 (0)	64	0,53
2	3 (0)	8 (3)	14 (9)	14 (5)	63	0,61
3	3 (1)	10 (8)	9 (8)	6 (2)	48	0,43

\* liczby w nawiasach oznaczają tzw. oczekiwane liczby zgodnych ocen, gdyby rozkład ocen był przypadkowy

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

Tab. 6. Powtarzalność oceny dla klasyfikacji Herringa  
Tab. 6. The intra-observer agreement for Catterall classification

Powtórzona ocena obserwatora	Zgodność każdej z grup wg Herringa			% zgodności	Współczynnik kappa
	A	B	C		
1	6 (1)*	22 (17)	11 (5)	67	0,53
2	8 (1)	33 (22)	13 (4)	87	0,80
3	3 (1)	29 (24)	9 (3)	71	0,48

\* liczby w nawiasach oznaczają tzw. oczekiwane liczby zgodnych ocen, gdyby rozkład ocen był przypadkowy

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy

Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy

Obserwator 3 – lekarz odbywający staż podyplomowy

Tab. 7. Powtarzalność oceny każdego z obserwatorów dla różnych grup wiekowych  
 Tab. 7. The intra-observer agreement for each of the observers for different age groups

Powtórzona ocena obserwatora	Klasyfikacja Catteralla				Klasyfikacja Herringa			
	% zgodności		współczynnik kappa		% zgodności		współczynnik kappa	
	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.	< 6 r.ż.	> 6 r.ż.
1	68	59	0,61	0,43	65	70	0,53	0,54
2	66	60	0,60	0,56	78	97	0,67	0,93
3	52	44	0,54	0,22	71	70	0,54	0,35

Obserwator 1 – specjalista II stopnia z 19-letnim stażem pracy  
 Obserwator 2 – specjalizujący się ortopeda z 5-letnim stażem pracy  
 Obserwator 3 – lekarz odbywający staż podyplomowy

badaczy poddawana jest krytyce ze względu na stosunkowo niski stopień powtarzalności i zgodności oceny.

Uzyskane przez nas wyniki zgodności między poszczególnymi parami obserwatorów (ang. inter – observer study) na poziomie zadowalającym, porównywalne są z wynikami uzyskanymi przez Specchiulliego i wsp. [4] oraz Ritterbuscha i wsp. [5], a nieco gorsze niż w pozostałych badaniach prowadzonych na ten temat [3,6,7,8]. Jak wynika z zestawień w Tabelach 1 i 2 najistotniejsze rozbieżności wystąpiły pomiędzy obserwatorami o największej różnicy w długości stażu pracy i doświadczenia zawodowego (obserwator 1 i 3), natomiast młodszy obserwatorzy oceniali stawy biodrowe „ostrzej”. Simmons i wsp. [8], w których pracy 15 obserwatorów oceniało 40 radiogramów, nie zauważyli takiej zależności, choć porównując wyniki oceny w poszczególnych podgrupach (rezydenci, lekarze w trakcie specjalizacji i tzw. konsultanci) zauważyli, że stopień zgodności oceny wzrasta wraz z doświadczeniem zawodowym.

Podobnie jak u Christensena i wsp. [3] oraz de Billy'ego i wsp. [6], również w naszych badaniach największa rozbieżność ocen dotyczy grupy II i III według klasyfikacji Catteralla. Christensen i wsp. [3] nawet po połączeniu grupy II i III w całość nie uzyskał lepszej zgodności oceny różnych obserwatorów, natomiast Hardcastle i wsp. [7] przeciwnie – przy łącznym rozpatrywaniu grupy II i III uzyskał dwukrotnie większą zgodność całości oceny (9 oceniających 69 radiogramów). Jednym z wniosków płynących z jego badań jest postulat uproszczenia klasyfikacji, znamienne, jeśli wziąć pod uwagę, że jednym z obserwatorów był sam jej twórca.

Pięć czynników ryzyka, zaproponowanych przez Catteralla jako uzupełnienie klasyfikacji, powoduje dalsze rozbieżności oceny. Billy i wsp. [6] oceniali

zgodność oceny występowania poszczególnych czynników ryzyka na materiale 23 radiogramów. Największą zgodność między dziewięcioma oceniającymi osobami zanotowano odnośnie występowania zęścia przynasady, najmniejszą w przypadku występowania ekstruzji głowy kości udowej z panewki. Wynikało to być może z faktu, iż w metodyce tego badania ekstruzję trzeba było określić jako obecną lub nieobecną, bez ściśle określonych kryteriów. Jest to zgodne z oryginalną pracą Catteralla [1], który występowanie tego czynnika ryzyka określa na podstawie poszerzenia przyśrodkowej części szczeliny stawowej, nie podając wartości bezwzględnych. Odnośnie pozostałych czynników ryzyka obserwatorzy w tych badaniach byli znacznie bardziej zgodni, nawet w przypadku objawu Gage'a, którego definicje autorstwa Catteralla [1] i samego Gage'a [9] są różne. Co ciekawe, wielu autorów poddaje w wątpliwość znaczenie prognostyczne tych właśnie pozostałych czterech czynników ryzyka [10,11,12].

Klasyfikację Herringa cechuje znacznie większa zgodność oceny między poszczególnymi obserwatorami, a uzyskany przez nas odsetek zgodności rzędu 76-79% oraz wartości współczynnika kappa rzędu 0,61 do 0,70 wskazują na dobrą i bardzo dobrą zgodność oceny. Wyniki te są porównywalne z wynikami innych [2,5,13,14], z wyjątkiem doniesienia Farsetti'ego i wsp. [15], który uzyskał zgodność rzędu 88%. Wyniku takiego nie udało się jak dotąd nikomu uzyskać. Zważywszy jednak jednoczesną 100% powtarzalność oceny przez tego samego obserwatora, taki wynik oceny może budzić pewne wątpliwości.

Z zebranych przez nas danych nie wynika już tak wyraźnie, jak w przypadku klasyfikacji Catteralla, jakoby młodszy stażem obserwatorzy oceniali radiogramy „ostrzej” oraz nie obserwuje się znacząco mniejszej zgodności oceny między obserwatorem najstarszym i najmłodszym. Do podobnych wniosków do-

szedł również Podeszwa i wsp. [14], który nie zaobserwował wpływu doświadczenia zawodowego oceniającego na stopień zgodności z innymi. Jego metoda rozszerzona była jednak o ocenę przedniej kolumny głowy kości udowej na radiogramach w projekcji Lauensteina, w sposób analogiczny, jak ocenia się kolumnę boczną w projekcji przednio-tylnej. Okazało się, że w prawie 1/5 przypadków oceniający klasyfikowali dane biodro o 1 stopień gorzej niż na radiogramach w projekcji a – p. Być może wynika to z faktu, iż proces martwiczy dotyka w tych przypadkach wyłącznie przedniej części głowy kości udowej, co nie do końca może być widoczne na radiogramach w projekcji przednio-tylnej.

Wyniki analizy powtarzalności oceny (ang. intra – observer study) według klasyfikacji Catteralla są na poziomie dobrym lub bardzo dobrym, z wyjątkiem wyników uzyskanych przez najmłodszego z obserwatorów. Odsetek zgodności rzędu 48% i współczynnik kappa 0,43 wskazują na fakt, iż klasyfikacja ta sprawia największe trudności osobom o mniejszym doświadczeniu. Wyniki powtarzalności oceny według klasyfikacji Herringa są dobre lub bardzo dobre, niezależnie od doświadczenia oceniającego, podobnie jak w opracowaniach Podeszwy i wsp. [14].

Po podziale materiału na dwie grupy wiekowe, poniżej i powyżej 6 roku życia, okazuje się, że klasyfikacja Herringa cechuje się większą zgodnością i powtarzalnością oceny radiogramów dzieci starszych. Klasyfikacja Catteralla jest bardziej powtarzalna dla dzieci młodszych niż dla starszych, choć otrzymaliśmy znacząco gorsze wyniki niż w klasyfikacji Herringa. I ponownie, wyniki uzyskane przez najmłodszego z oceniających nie osiągają w przypadku klasyfikacji Catteralla zadowalającego poziomu zgodności i powtarzalności w przypadku radiogramów dzieci starszych.

Porównując wyniki analizy statystycznej zgodności i powtarzalności oceny według każdego z ww. systemów klasyfikacyjnych, widoczne są poza tym wyższe wartości współczynnika kappa dla klasyfikacji Catteralla w przypadku powtarzania oceny przez tego samego obserwatora niż w przypadku zgodności oceny z innym obserwatorem [15]. Może to świadczyć o subiektywnym charakterze tej klasyfikacji.

## WNIOSKI

1. Klasyfikacja Catteralla cechuje się mniejszą powtarzalnością i zgodnością oceny niż klasyfikacja Herringa, zwłaszcza w odniesieniu do dzieci powyżej 6 roku życia.
2. Największe trudności oceniającym sprawia rozróżnienie grupy II i III w klasyfikacji Catteralla.

## PIŚMIENNICTWO

1. Catterall A: The natural history of Perthes' disease. *J Bone Joint Surg (Br)* 1971; 53-B: 37-53.
2. Herring JA, Neustadt JB, Williams JJ, Early JS, Browne RH: The lateral pillar classification of Legg-Calvé-Perthes disease. *J Pediatr Orthop* 1992; 12: 143-150.
3. Christensen F, Soballe K, Ejsted R, Luxhoj T: The Catterall classification of Perthes disease. An assessment of reliability. *J Bone Joint Surg (Br)* 1986; 68-B: 614-615.
4. Specchiulli F, Scialpi L: La classificazione di Catterall versus la classificazione di Herring nella malattia di Perthes. *Chir Organi Mov* 1997; 82: 289-293.
5. Ritterbusch JF, Shantharam SS, Gelinas C: Comparison of lateral pillar classification and Catterall classification of Legg-Calvé-Perthes' disease. *J Pediatr Orthop* 1993; 13: 200-202.
6. de Billy B, Viel JF, Monnet E, Garnier E, Aubert D: Interobserver reliability in the interpretation of radiologic signs in Legg – Calvé – Perthes disease. *J Pediatr Orthop Part B* 2002; 11: 10-14.
7. Hardcastle PH, Ross R, Hamalainen M, Mata A: Catterall grouping of Perthes disease. An assessment of observer error and prognosis using the Catterall classification. *J Bone Joint Surg Br* 1980; 62-B: 428-431.
8. Simmons ED, Graham HK, Szalai JP: Interobserver variability in grading Perthes' disease. *J Bone. Joint Surg Br* 1990; 72-B: 202-4.
9. Gage HC: Possible early sign of Perthes' disease. *Br J Radiol* 1933; 6: 295-297.
10. Ismail AM, Macnicol MF: Prognosis in Perthes' disease. A comparison of radiological predictors. *J Bone Joint Surg (Br)* 1998; 80-B: 310-314.
11. Klisić PJ: Perthes' disease. *Int Orthop* 1984; 8: 95-102.
12. Mukherjee A, Orth D, Fabry G: Evaluation of the prognostic indices in Legg-Calvé-Perthes disease: Statistical analysis of 116 hips. *J Pediatr Orthop* 1990; 10: 153-158.
13. Lappin K, Kealey D, Cosgrove A: Herring classification: How useful is the initial radiograph? *J Pediatr Orthop* 2002; 22: 479-482.
14. Podeszwa DA, Stanitski CL, Stanitski DF, Woo R, Mendelow MJ: The effect of pediatric orthopaedic experience on interobserver and intraobserver reliability of the Herring lateral pillar classification of Perthes disease. *J Pediatr Orthop* 2000; 20: 562-565.
15. Farsetti P, Tudisco C, Caterini R, Potenza V, Ippolito E: The Herring lateral pillar classification for prognosis in Perthes disease. Late results in 49 patients treated conservatively. *J Bone Joint Surg (Br)* 1995; 77-B: 739-742.
16. Pietrzak S, Napiontek M, Tomaszewski M.: Zgodność klasyfikacji radiologicznych Catteralla i Herringa w ocenie choroby Perthesa. *Chir Narz Ruchu Ortop Pol* 2000; 65: 33-38.

*Adres do korespondencji / Address for correspondence*  
Szymon Pietrzak  
Klinika Ortopedii CMKP  
05-400 Otwock, ul. Konarskiego 13  
szymon.pietrzak@inetia.pl

Otrzymano / Received 23.07.2004 r.  
Zaakceptowano / Accepted 13.09.2004 r.