Clin Orthop Relat Res (2013) 471:2259–2265
DOI 10.1007/s11999-013-2844-7

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CLINICAL RESEARCH

Validity and Reliability of the Paprosky Acetabular Defect Classification

Raymond Yu BM, BS, HB.MedSc, M.SurgSci, Jochen G. Hofstaetter MD, Thomas Sullivan BMa&CompSc(Hons), Kerry Costi BA, Donald W. Howie BM, BS, PhD, FRACS, Lucian B. Solomon MD, PhD, FRACS

Received: 22 September 2012 / Accepted: 31 January 2013 / Published online: 15 February 2013

Abstract

Background The Paprosky acetabular defect classification is widely used but has not been appropriately validated. Reliability of the Paprosky system has not been evaluated in combination with standardized techniques of measurement and scoring.

Questions/purposes This study evaluated the reliability, teachability, and validity of the Paprosky acetabular defect classification.

Methods Preoperative radiographs from a random sample of 83 patients undergoing 85 acetabular revisions were classified by four observers, and their classifications were compared with quantitative intraoperative measurements. Teachability of the classification scheme was tested by dividing the four observers into two groups. The observers in Group 1 underwent three teaching sessions; those in Group 2 underwent one session and the influence of teaching on the accuracy of their classifications was ascertained.

Results Radiographic evaluation showed statistically significant relationships with intraoperative measurements of anterior, medial, and superior acetabular defect sizes. Interobserver reliability improved substantially after teaching and did not improve without it. The weighted kappa coefficient went from 0.56 at Occasion 1 to 0.79 after three teaching sessions in Group 1 observers, and from 0.49 to 0.65 after one teaching session in Group 2 observers.

Conclusions The Paprosky system is valid and shows good reliability when combined with standardized definitions of radiographic landmarks and a structured analysis.

Level of Evidence Level II, diagnostic study. See the Guidelines for Authors for a complete description of levels of evidence.

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request.
Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.
This work was performed at the Centre for Orthopaedic and Trauma Research, University of Adelaide, Adelaide, Australia.

R. Yu, J. G. Hofstaetter, K. Costi, D. W. Howie, L. B. Solomon (✉)
Orthopaedic & Trauma Service, Department of Orthopaedic & Trauma, Royal Adelaide Hospital, Level 4 Bice Building, North Terrace, Adelaide, SA 5000, Australia
e-mail: bogdan.solomon@health.sa.gov.au

J. G. Hofstaetter, D. W. Howie, L. B. Solomon
Centre for Orthopaedic and Trauma Research, University of Adelaide, Adelaide, Australia

J. G. Hofstaetter
Department of Orthopaedics, Vienna General Hospital, Medical University of Vienna, Vienna, Austria

T. Sullivan
Discipline of Public Health, University of Adelaide, Adelaide, Australia

Introduction

Revision THA is being performed with increasing frequency [1]. With a younger cohort of patients and the increased lifespan of patients living with a hip prosthesis, the absolute number of failed prostheses and revision

procedures is likely to continue rising. Preoperative planning is an essential part of revision THA and preoperative pelvic radiographs are most commonly used for assessment of periprosthetic bone stock and defects. Multiple classification systems have been devised to describe the acetabular bone defects before revision THA: Paprosky [15], D'Antonio [3], or American Academy of Orthopaedic Surgeons (AAOS); Saleh [18], Gustilo [9], Gross [8], Parry [16], and Engh [5, 10]. Classifications are important not only to guide for the revision technique, but also to render the results in different clinical studies comparable.

The Paprosky system, first described in 1994 [14], may be the most widely used acetabular defect classification system. This system classifies defects according to the presence or absence of intact acetabular walls, and the ability of the anterior and posterior columns to support an implant [14, 17, 19]. Type 1 defects have only minimal bone loss, no component migration, and intact acetabular walls. Type 2 defects have moderate bone loss with distortion of the acetabular hemisphere but preservation of the anterior and posterior acetabular columns. The destruction involves the superior and/or the medial walls and based on its location, Type 2 defects are subclassified into 2A (direct superior), 2B (superolateral), and 2C (medial). Type 3 defects show severe bone loss from major destruction of the acetabular rim and supporting structures; these are subclassified into 3A and 3B. Type 3A defects include moderate destruction of the medial wall and posterior column, while Type 3B defects show complete destruction of the medial wall and severe destruction of the posterior column [15]. Based on the Paprosky acetabular defect type, recommendations can be made for grafting and fixation strategies [13].

Despite its popularity, there remains dispute over the reliability and validity of the Paprosky system [2, 7, 14]. In general, for a classification system to be effective in classifying acetabular defects and in guiding the revision technique and assisting outcome comparisons, it must possess a high degree of reliability and validity. Reliability refers to the reproducibility of a classification system within one and among several observers, whereas validity refers to the ability for the system to accurately describe or predict the true pathology when compared with a reference standard [6]. The validity of the Paprosky system would be best assessed by comparing it with the gold standard of intraoperative findings.

To date, no study has validated the Paprosky classification system using quantitative intraoperative measurements of bone loss. Furthermore, no study has evaluated whether consistency in definition and interpretation of the radiographic landmarks used in the Paprosky system and a structured analysis can improve its interobserver and intraobserver reliability.

The aims of this study were (1) to determine whether the preoperative classification of acetabular bone deficiency using the Paprosky system correlates with intraoperative measurements of bone deficiency; and (2) to investigate whether the reliability of the system improves with formal teaching of the classification, with explicit standardization of the interpretation of radiographic landmarks and a structured analysis of all the elements of the classification.

Patients and Methods

Data were obtained from patients who underwent revision THAs performed at a large orthopaedic unit at the Royal Adelaide Hospital (Adelaide, South Australia) from January 2000 to April 2010. This study and protocol were approved by the local institution's research ethics committee.

We included patients in the study if an acetabular component was revised and if preoperative AP pelvic radiographs and intraoperative records were complete. As per the description of the Paprosky classification system [15], only AP pelvic radiographs were used to classify the acetabular defects investigated in this study. Patients were excluded from the study if revision THA was performed in the absence of radiographic osteolysis, that is, where the primary indication for revision was for pain, recurrent dislocation, or infection without radiographic osteolysis present. We also excluded patients if substantial intraoperative acetabular fractures occurred, or if preoperative pelvic radiographs were of insufficient quality. For example, pelvic radiographs were excluded if they were not true AP views. We also collected patient characteristics, including age, sex, date of revision, implant type removed, and implant type inserted.

A simulation study indicated that a sample size of 85 would allow for the estimation of interobserver or intraobserver reliability with precision $\pm$ 0.18 or better, in which precision refers to the width of a 95% CI around a weighted kappa statistic. The calculations assume the population-weighted kappa is unknown but is larger than 0.25, which corresponds to weak agreement. The calculations also assume the population proportion of patients in each of the six classifications of the Paprosky scale is the same as described in the study by Campbell et al. [2].

We performed a total of 249 revision THAs involving revision of an acetabular component during the study period. Twenty-five cases were excluded because no digital copies of preoperative radiographs were available, 13 were excluded because the preoperative radiographs were deemed to be of insufficient quality, and 31 were excluded as a result of incomplete intraoperative data. This left a total of 180 radiographs that met our inclusion criteria available for review. These were numbered consecutively

and a random number generator (85 numbers, from 1 to 180) then was used to select 85 for inclusion in the study by an independent observer (SK). Fourteen procedures were performed for the primary indication of late recurrent dislocation, 48 were performed for prosthetic loosening, and 23 for osteolysis.

Acetabular defect dimensions were recorded intraoperatively by six operating surgeons (DWH, LBS, RC, AM, DT, SB) with 75 being recorded by two surgeons (DWH, LBS). After removal of the primary implant, the length and width of the defects at the lateral, medial, anterior, and posterior acetabular walls were measured using rulers and recorded in centimeters to the nearest 0.5 cm. The dimensions of acetabular defects measured intraoperatively were recorded (Fig. 1). Defect type, either cortical or cavitary, also was recorded. Intraoperative data, including an acetabular diagram with a scaled ruler, were recorded on a standardized operation record form. Data were entered and extracted from the department's Orthopaedic Patient Management and Outcomes Documentation Database. Defect surface area at each acetabular wall was estimated in cm^2.

A guide for scoring the radiographs was developed by the authors based on the Paprosky system (Table 1) [15]. This guide scores four radiographic features identified by the Paprosky system: (1) teardrop lysis/defect; (2) ischial lysis/defect; (3) the ilioischial (Kohler's) line; and (4) superior migration of the hip center/superior defect. The degree of teardrop defects is grouped into three grades depending on the presence of the lateral and medial teardrop borders. Ischial defects are graded depending on the measured size of ischial osteolytic lesions. Kohler's line is graded depending on its disruption. Superior defects are grouped into four grades depending on the vertical length of the defect from the original center of rotation of the femoral head and the area of acetabular dome involvement.

We used a second scoring table to assign the overall Paprosky type (Table 2). This was done to ensure a structured scoring mechanism in which all scoring elements are considered and isolated defects automatically lead to a worse score. On the occasions in which combinations of radiographic features for a given case did not meet all the criteria for one Paprosky type, or by contrary, met the criteria for more than one Paprosky type, the least

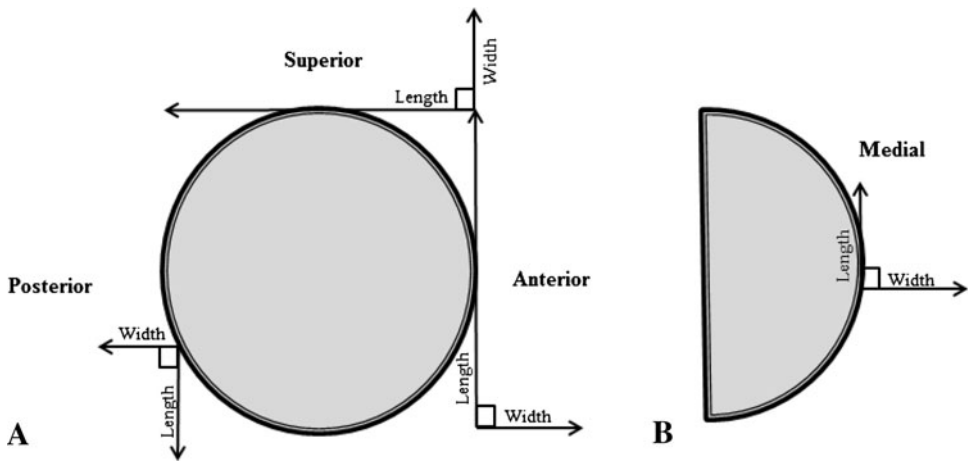

Fig. 1A–B (**A**) Lateral and (**B**) AP views of the acetabulum show the dimensions measured intraoperatively at the superior, posterior, anterior, and medial walls.

Table 1. Scoring guide for assessment of the radiographic features of the Paprosky system

Radiographic feature	Feature grading			
Teardrop defect	T1		T2	T3
	Intact		Moderate	Severe
			(loss of lateral border)	(loss of medial border)
Ischial defect	I1	I2	I3	I4
	None	Mild	Moderate	Severe
		(< 0.7 cm)	(0.7–1.4 cm)	(> 1.4 cm)
Kohler's line	K1		K2	
	Intact		Disrupted	
Superior defect	S1	S2	S3	S4
	Minimal	Mild	Moderate	Severe
		(< 3 cm superior)	(< 3 cm superolateral)	(> 3 cm)

Table 2. Scoring guide for overall Paprosky type based on the four radiographic features

Teardrop defect	Ischial defect	Kohler's line	Superior defect	Paprosky type
T1	I1	K1	S1	I
T1	I2	K1	S2	IIA
T1	I2	K1	S3	IIB
T2	I2	K2	S2, S3	IIC
T2	I3	K1	S4	IIIA
T3	I4	K2	S4	IIIB

severe Paprosky type and the one with most radiographic features met were selected.

Radiographs were analyzed on three separate occasions by four observers (LBS, RY, JGH, MP), three of whom were trained consultant-level orthopaedic surgeons specializing in hip arthroplasty and the other a recent medical graduate (junior medical officer) working in the orthopaedic unit. Each observation occasion was separated by at least 2 weeks as a washout period before rerandomization of radiographs. Before each occasion of radiographic analysis, two of the observers, an orthopaedic surgeon and a junior medical officer (Group 1), underwent some teaching sessions together using the scoring guide to ensure standardized interpretation of radiographic features. Teaching involved analysis and discussion of 10 random cases other than the ones being investigated to ensure agreement on definitions of radiographic landmarks when using the scoring table and to resolve discrepancies in the interpretation of the Paprosky system types. The remaining two observers, both of whom were orthopaedic surgeons (Group 2), analyzed the radiographs on two occasions without undergoing any prior teaching session and without using the scoring system described. Before the third radiograph analysis, these investigators underwent a similar teaching session as Group 1.

We compared intraoperative data across Paprosky classifications using nonparametric Kruskal-Wallis tests, comparing anterior defect area among the three teardrop classification groups, posterior defect area among the four ischium classification groups, medial defect area between the two Kohler's line classification groups, and superior defect area among the four superior defect classification groups. Results with a p value less than 0.05 were considered statistically significant.

We determined interobserver and intraobserver reliability using weighted kappa coefficients. Differences in interobserver reliabilities with time were assessed using the nonparametric bootstrapping method based on 10,000 bootstrap samples. Extent of agreement was interpreted using the criteria described by Landis and Koch [11], such that a score greater than 0.80 indicates excellent agreement, a score ranging from 0.61 to 0.80 indicates good agreement, a score of 0.41 to 0.60 indicates moderate agreement, a score of 0.21 to 0.40 indicates fair agreement, and a score of 0.20 or less indicates poor agreement. Radiographs were reidentified by a third party (SK) and intraoperative data were correlated against corresponding preoperative Paprosky types. All statistical calculations were performed using SAS Version 9.3 (SAS Institute Inc, Cary, NC, USA).

Table 3. Prevalence of Paprosky defect type (n = 85)

Paprosky type	Number	Prevalence
Type 1	23	27%
Type 2A	25	29%
Type 2B	14	17%
Type 2C	7	8%
Type 3A	7	8%
Type 3B	9	11%

In this series, Type 2A defects were the most prevalent defect type, seen in 25 (29%) of the cases analyzed, followed by Type 1 defects comprising 23 (27%) cases. Nine (11%) analyzed cases were classified as the most severe Type 3B defects (Table 3). These data were derived from analyses performed after three teaching sessions by an experienced, consultant, orthopaedic surgeon.

Results

When we compared defect surface area at each acetabular wall with its corresponding radiographic feature, we found numerous important relationships. A more severe radiologic teardrop defect was associated with a larger intraoperative defect area ($p = 0.0015$) (Fig. 2). A more severe Kohler's line grade was associated with a larger intraoperative defect area ($p = 0.0011$) (Fig. 3). A more severe superior radiographic defect was associated with a larger intraoperative superior defect area ($p < 0.0001$) (Fig. 4). There was no evidence of a relationship in the distribution of intraoperative posterior defect area and the four ischium classification groups overall ($p = 0.21$), but large intraoperative ischial defects were assessed as being severe on preoperative radiographs (Fig. 5).

In Group 1 (the observers who received three teaching sessions), the teaching sessions improved reliability between the first and the third sessions (0.56 versus 0.79, $p < 0.001$). Interobserver reliability did not improve significantly between Occasion 1 and Occasion 2 (weighted kappa 0.56 versus 0.69, $p = 0.071$) or between Occasion 2 and Occasion 3 (0.69 versus 0.79, $p = 0.052$). Using the criteria of Landis and Koch [11], interobserver agreement between the Group 1 observers improved from moderate to

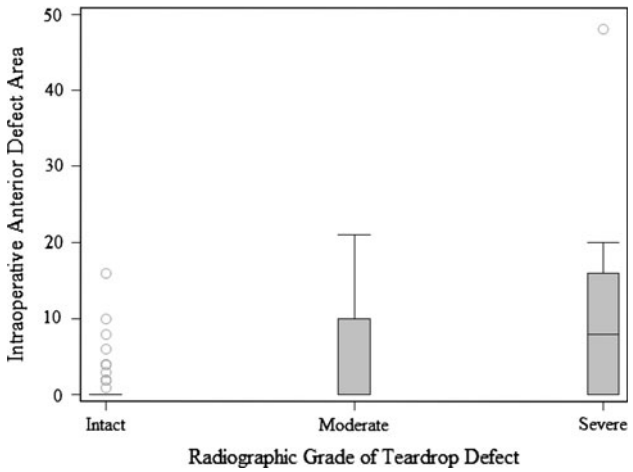

Fig. 2 The box and whisker plot shows the surface area of an anterior defect measured intraoperatively compared with a teardrop defect according to the Paprosky radiographic classification.

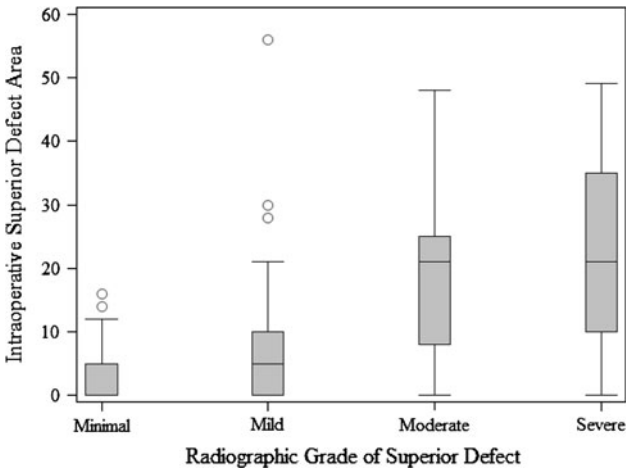

Fig. 4 This box and whisker plot shows the surface area of a superior defect measured intraoperatively compared with a superior acetabular defect according to the Paprosky radiographic classification.

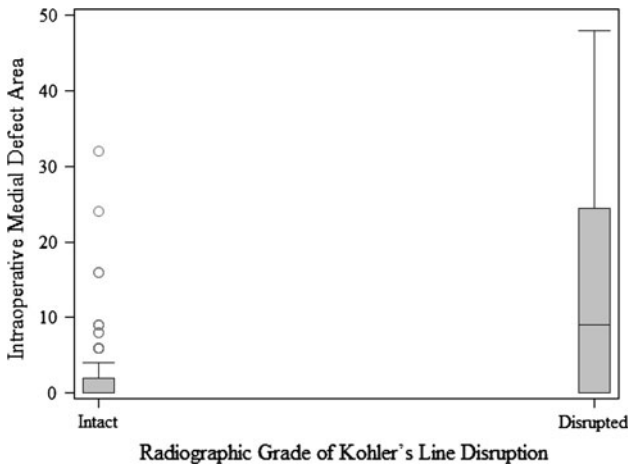

Fig. 3 The box and whisker plot shows the surface area of a medial defect measured intraoperatively compared with Kohler's line according to the Paprosky radiographic classification.

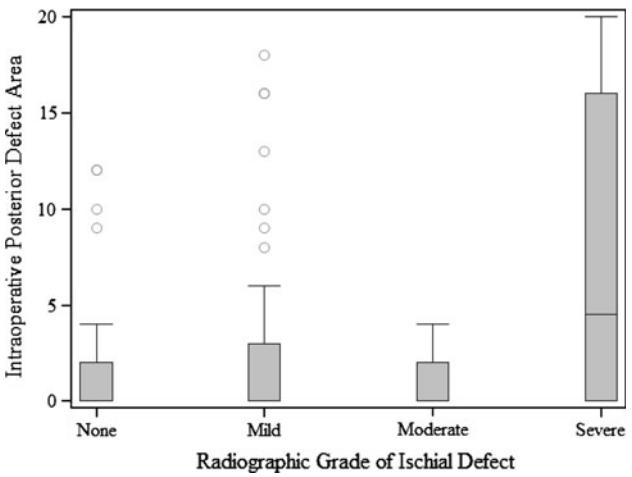

Fig. 5 A box and whisker plot shows the surface area of a posterior defect measured intraoperatively compared with an ischial defect according to the Paprosky radiographic classification.

good between Occasions 1 and 3. Intraobserver agreement was calculated for both observers in Group 1 between Occasions 1 and 3, resulting in kappa = 0.66 (95% CI, 0.56–0.76) for Observer 1 and kappa = 0.71 (95% CI, 0.62–0.81) for Observer 2. These values would be considered good agreement.

Group 2 observers did not undergo standardized teaching sessions before Occasions 1 and 2, but instead were asked to base their radiographic assessments on their individual interpretations of the Paprosky classification system. Interobserver agreement between these two observers on Occasions 1 and 2 were kappa = 0.49 (95% CI, 0.38–0.61) and kappa = 0.50 (95% CI, 0.43–0.64), respectively. This is equivalent to moderate agreement on both occasions with no significant improvement in reliability between each occasion ($p = 0.60$). After a washout period and a single teaching session with Group 1 observers, Group 2 observers underwent a third period of observation. Interobserver reliability between the Group 2 observers on the third occasion improved significantly to good agreement with kappa = 0.65 ($p = 0.014$).

Discussion

The appropriate preoperative acetabular defect classification system remains debatable [2, 4, 10, 12]. A valid and reliable preoperative classification system can assist in effective preoperative planning and allow comparisons to be made between different revision techniques [6]. We therefore examined whether the preoperative classification

of acetabular bone deficiency using the Paprosky system correlates with intraoperative measurements of bone deficiency and if the reliability of the system improves with formal teaching, which included standardization of the interpretation of radiographic landmarks, and a structured analysis of all the elements of the classification.

This study has several limitations. There is a limitation to the degree of information gleaned from a comparison between categorical and quantitative data; however, it is evident that higher-grade radiographic defects according to the Paprosky system are associated with greater surface area defects at the anterior, medial, and superior acetabular walls. Another limitation is that the intraoperative defect measurements were obtained by more than one surgeon during the study's 10-year period in which patients were included. A standardized measurement technique and rigor cannot be ensured despite the use of standard intraoperative forms including acetabular diagrams to be filled out by the surgeon. In addition, intraoperative measurements were rounded off to the nearest 0.5 cm when measurements to the nearest millimeter would be ideal given that radiographic grades for ischial defects were separated by 0.7-cm intervals. Moreover, intraoperative defects were recorded as surface areas and calculated as two-dimensional rectangles, when in reality defects possessed irregular shapes of varying volumes. Calculated defect areas therefore were only approximations of true defect proportions.

In this study, we validated the Paprosky acetabular defect system using quantitative intraoperative measurements of the defects. We observed a significant correlation between defects at the anterior, medial, and superior acetabular walls and their respective radiographic landmarks. No relationship was seen between posterior acetabular defect size and ischial defect grade noted on radiographs. This is likely because the Paprosky classification is based on AP pelvis radiographs and in this view, the posterior acetabulum often is obscured by a radiopaque acetabular prosthesis. Thus, less importance should be placed on evaluation of the ischium on AP views and oblique and lateral hip views should be used to assess the posterior acetabulum. Despite this issue, large intraoperative ischial defects were still assessed as being severe on preoperative radiographs.

In this study, we also assessed the impact of teaching on the reliability of the Paprosky system. We observed an improvement in interobserver reliability after teaching sessions in which interpretation of radiographic features was discussed between observers. Standardization was further ensured with the use of a structured scoring table as a guide. Observers from Group 1 underwent teaching sessions before each observation period. Subsequently, the interobserver reliability attained after each occasion improved significantly, and after the third occasion was higher than any previously published values for interobserver reliability for the Paprosky system [2, 7, 16] and approached a score of excellent. Intraobserver reliability was similar for both observers, excluding the possibility that one observer simply tended toward the other after the second teaching session. Group 2 observers, composed of two consultant-level orthopaedic surgeons, showed that before undergoing a teaching session to standardize the interpretation of radiographic classification, reliability of the Paprosky system was only moderate and comparable to previously published reliability scores [2, 7]. After a single teaching session, interobserver reliability between these observers, and also between Group 1 observers, improved to good strength.

Our study further compared preoperative Paprosky classification with corresponding intraoperative acetabular defect measurements. There are few data analyzing standardized quantitative intraoperative measurements of acetabular defects, yet this should be the gold standard required to validate such a preoperative classification system. Commonly used preoperative classifications and the operative plans derived from these either may be validated or rejected. In the event of system validation, different revision and grafting techniques from multiple centers and series can be more objectively evaluated. Gozzard et al. investigated the reliability of the Paprosky classification and also examined validity by drawing intraoperative comparisons [7]. Interestingly, despite poor reliability, good validity was found. However, their study was small, analyzing only 25 patients undergoing revision THAs, and did not describe a standardized and quantitative system for measuring intraoperative findings. Campbell et al. compared the interobserver and intraobserver reliabilities for the Paprosky, AAOS, and Gross classification systems using 33 preoperative radiographs and three groups of observers of different expertise [2]. Their study revealed that even among the originators of each respective classification system, intraobserver reliability was underwhelming, and poor overall interobserver reliability was shown. These findings were echoed by Gozzard et al. in an analysis of the AAOS and Paprosky systems [7]. Neither of these studies commented on the teachability of the Paprosky system despite the likely improvement that teaching would have on standardizing the interpretation of radiographic features.

Our findings suggest the Paprosky acetabular classification system is valid, but can be subjective and requires a standardized measurement technique and standardized method of scoring to decrease interobserver differences. In contrast to some previous studies, our study showed that teaching sessions combined with the use of a scoring guide substantially improves reliability, from moderate agreement to good agreement, and we recommend that future studies publishing on this system describe steps they took to achieve standardization in the classification's use.

Acknowledgments We thank Mark Price, BM, BS, FRCS, for being one of the observers to grade the radiographs and Serena Khoo for helping with deidentification and randomization of radiographs and reviewing and editing the manuscript. We also acknowledge Richard Clarnette BM, BS, FRACS, Andrew Mintz BM, BS, FRACS, Daryl Teague BM, BS, FRACS, and Scott Brumby BM, BS, PhD, FRACS, for providing patient data.

References

1. Australian Orthopaedic Association National Joint Registry. Annual Report. Adelaide, Australia: The Australian Orthopaedic Association; 2010.
2. Campbell DG, Garbuz DS, Masri BA, Duncan CP. Reliability of acetabular bone defect classification systems in revision total hip arthroplasty. *J Arthroplasty.* 2001;16:83–86.
3. D'Antonio JA. Periprosthetic bone loss of the acetabulum: classification and management. *Orthop Clin North Am.* 1992;23:279–290.
4. Davis AM, Schemitsch EH, Gollish JD, Saleh KJ, Davey R, Kreder HJ, Mahomed NN, Waddell JP, Szalai JP, Gross AE. Classifying failed hip arthroplasty: generalizability of reliability and validity. *Clin Orthop Relat Res.* 2003;415:171–179.
5. Engh CA, Glassman AH. Cementless revision of failed total hip replacement: an update. *Instr Course Lect.* 1991;40:189–197.
6. Garbuz DS, Masri BA, Esdaile J, Duncan CP. Classification systems in orthopaedics. *J Am Acad Orthop Surg.* 2002;10:290–297.
7. Gozzard C, Blom A, Taylor A, Smith E, Learmonth I. A comparison of the reliability and validity of bone stock loss classification systems used for revision hip surgery. *J Arthroplasty.* 2003;18:638–642.
8. Gross AE, Duncan CP, Garbuz D, Mohamed EM. Revision arthroplasty of the acetabulum in association with loss of bone stock. *Instr Course Lect.* 1999;48:57–66.
9. Gustilo RB, Pasternak HS. Revision total hip arthroplasty with titanium ingrowth prosthesis and bone grafting for failed cemented femoral component loosening. *Clin Orthop Relat Res.* 1988;235:111–119.
10. Johanson NA, Driftmier KR, Cerynik DL, Stehman CC. Grading acetabular defects: the need for a universal and valid system. *J Arthroplasty.* 2010;25:425–431.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
12. Masri BA, Masterson EL, Duncan CP. The classification and radiographic evaluation of bone loss in revision hip arthroplasty. *Orthop Clin North Am.* 1998;29:219–227.
13. O'Rourke MR, Paprosky WG, Rosenberg AG. Use of structural allografts in acetabular revision surgery. *Clin Orthop Relat Res.* 2004;420:113–121.
14. Paprosky WG, Bradford MS, Younger TI. Classification of bone defects in failed prostheses. *Chir Organi Mov.* 1994;79:285–291.
15. Paprosky WG, Perona PG, Lawrence JM. Acetabular defect classification and surgical reconstruction in revision arthroplasty. A 6-year follow-up evaluation. *J Arthroplasty.* 1994;9:33–44.
16. Parry MC, Whitehouse MR, Mehendale SA, Smith LK, Webb JC, Spencer RF, Blom AW. A comparison of the validity and reliability of established bone stock loss classification systems and the proposal of a novel classification system. *Hip Int.* 2010;20:50–55.
17. Pluot E, Davis ET, Revell M, Davies AM, James SL. Hip arthroplasty. Part 2: normal and abnormal radiographic findings. *Clin Radiol.* 2009;64:961–971.
18. Saleh KJ, Holtzman J, Gafni A, Saleh L, Jaroszynski G, Wong P, Woodgate I, Davis A, Gross AE. Development, test reliability and validation of a classification for revision hip arthroplasty. *J Orthop Res.* 2001;19:50–56.
19. Sporer SM, Paprosky WG, O'Rourke MR. Managing bone loss in acetabular revision. *Instr Course Lect.* 2006;55:287–297.