Utility of Radiographs, Computed Tomography, and Three Dimensional Computed Tomography Pelvis Reconstruction for Identification of Acetabular Defects in Residency Training

Johannes F. Plate, MD, PhD*, John S. Shields, MD*, Maxwell K. Langfitt, MD*,
Michael P. Bolognesi, MD†, Jason E. Lang, MD‡, Thorsten M. Seyler, MD, PhD*†

*Department of Orthopaedic Surgery, Wake Forest School of Medicine, Winston-Salem, NC, USA**
Department of Orthopaedic Surgery, Duke University, Durham, NC, USA†
Blue Ridge Bone and Joint, Arden, NC, USA‡

Purpose: The Paprosky classification system of acetabular defects is complex and its reliability has been questioned. The purpose of this study was to evaluate the effectiveness of different radiologic imaging modalities in classifying acetabular defects in revision total hip arthroplasty (THA) and their value of at different levels of training.
Materials and Methods: Bone defects in 8 revision THAs were classified by 2 fellowship-trained adult reconstruction surgeons. A timed presentation with representative images for each case (X-ray, two-dimensional computed tomography [CT] and three-dimensional [3D] reconstructions) was shown to 35 residents from the first postgraduate year of training year of training (PGY-1 to PGY-5), 2 adult reconstruction fellows and 2 attending orthopaedic surgeons. The Paprosky classification of bone defects was recorded. The influence of image modality and level of training on classification were analyzed using chi-square analysis (alpha=0.05).
Results: Overall correct classification was 30%. The level of training had no influence on correct classification (P=0.531). Using X-ray led to 37% correctly identified defects, CT scans to 33% and 3D reconstructions to 20% of correct answers (P<0.001). There was no difference in correct classification based defect type (P<0.001). Regardless of level of training or imaging, 64% of observers recognized type 1 defects, compared to only 16% correct recognition of type 3B defects.
Conclusion: Using plain X-rays led to an increased number of correct classification, while regular CT scan and 3D CT reconstructions did not improve accuracy. The classification system of acetabular defects can be used for treatment decisions; however, advanced imaging may not improve its utilization.

Key Words: Hip replacement arthroplasty, Acetabular defect, Classification, Reoperation

INTRODUCTION

The number of complex revision total hip arthroplasties (THA) is predicted to rise. The identification of acetabular bone defects prior to revision THA has important implications on technique and complexity of acetabular reconstruction. A majority of patient presenting for revision THA has some degree of acetabular bone loss[1]. Several classification systems for acetabular bone loss have been described[1-4]; however, the Paprosky classification system is most widely used. The Paprosky classification system includes three main types with up to three subtypes focused on the integrity of the superior rim of the acetabulum and medial wall. This system also provides a practical application to guide the surgeon to the appropriate treatment option. However, the classification system is complex and its intraobserver and interobserver reliability has been questioned[4].

The original description of the classification system used only anterior-posterior pelvic X-rays for evaluation of bone defects[1]. The addition of multiplanar computed tomography (CT) as an adjunct to plain X-rays was found to increase the sensitivity of detecting bone loss[5]. Recently, three-dimensional (3D) reconstructed CT scans have become widely available potentially increasing the ability to identify acetabular bone loss[6]. The use of 3D imaging was found to improve the familiarity of the complex spatial anatomy of the acetabulum and was shown to improve the recognition of acetabular fractures[7]. However, the additional use of 3D images for classifying acetabular defects using the Paprosky classification has not been evaluated.

The purpose of this study was to evaluate the effectiveness of different radiologic imaging modalities (radiographs, standard multiplanar CT, 3D CT reconstructions) in classifying acetabular defects in revision hip arthroplasty cases and their value of at different levels of orthopaedic training. The hypothesis was that 3D CT reconstruction enhances the ability to classify acetabular defects correctly.
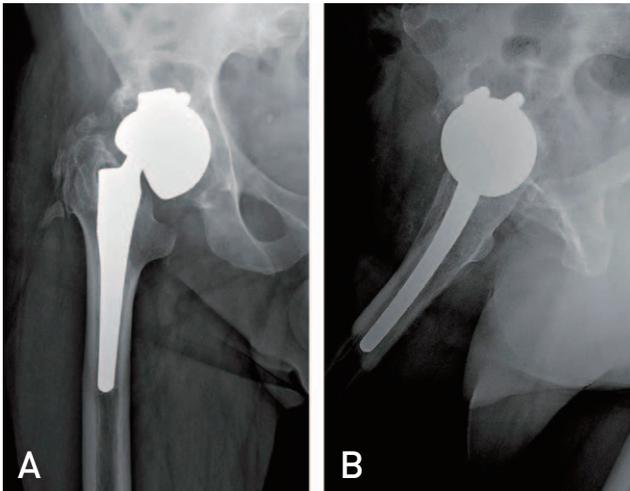
MATERIALS AND METHODS

All patients treated with revision THA with acetabular bone defects between 2002 and 2012 were identified. Inclusion criteria were patients that had plain X-rays, standard multiplanar CT scan and 3D CT reconstructions representing each Paprosky type defect available for review. Bone defects were classified independently by two fellowship-trained adult reconstruction surgeons (one type 1 defect, one type 2A, two type 2B, two type 2C, one type 3A, and one type 3B acetabular defects) based on the Paprosky classification of acetabular bone loss (Table 1)[1,8]. X-rays, CT scan and 3D CT reconstructions were classified based on several anatomic landmarks described in the classification system. These landmarks included the presence or absence of the teardrop representing the medial wall, migration of the hip center representing bone loss of the superior acetabular dome, continuity of Kohler's line representing the anterior column, involvement of the ischium or posterior column and overall amount of bone loss (Table 1)[1,8]. All observers received a formal educational lecture on how to apply the classification system and what landmarks to assess prior the timed presentation of the cases. Representative sections of the multiplanar CT scans were chosen by the examiners and compiled into a timed presentation. The presentation consisted of plain X-rays (Fig. 1), selected sections of multiplanar CT scans (Fig. 2), and 3D reconstructions (Fig. 3) for each defect type on separate slides and in random order similar to previous study[7]. 3D reconstructions of one type 2C

Table 1. Critera for Assessment of the Paprosky Classification System for Acetabular Bone Defects

| Type	Femoral head center migration	Kohler's line	Teardrop	Ischial osteolysis
I	None	Intact	Intact	None
IIA	Mild (<2 cm superomedial)	Intact	Intact	None
IIB	Moderate (<2 cm superolateral)	Intact	Intact	Mild
IIC	Mild (<2 cm medial)	Disrupted	Moderate lysis	Mild
IIIA	Severe (>2 cm superolateral)	Intact	Moderate lysis	Moderate
IIIB	Severe (>2 cm superomedial)	Disrupted	Severe lysis	Severe

Adapted from the article of Telleria and Gee (Clin Orthop Relat Res 2013;471:3725-30)[8].

Hip & Pelvis

Johannes F. Plate et al. Imaging Modalities to Identify Acetabular Defects in Residency Training

Fig. 1. Anterior-posterior (**A**) and lateral (**B**) radiograph of a Paprosky type 2A acetabular defect.

defect were excluded due to the low quality of the outside CT scan subsequent increased metal artifact.

Thirty-five residents from the first postgraduate year of training (PGY-1) to the final 5th year of training (PGY-5), 2 adult reconstruction fellows and 2 attending orthopaedic surgeons were recruited for this study and received a 15-minute introduction to the classification system. The 22 slides were then shown for 60 seconds in random order of defect type and imaging modality and participants were asked to independently identify the defect type. Chi-square analysis was utilized to examine the influence of image modality and level of training on the correct classification of acetabular bone loss using the Paprosky classification system with alpha=0.05. Krippendorff's alpha coefficient (α) was calculated to assess the overall interobserver reliability of the Paprosky classification for imaging
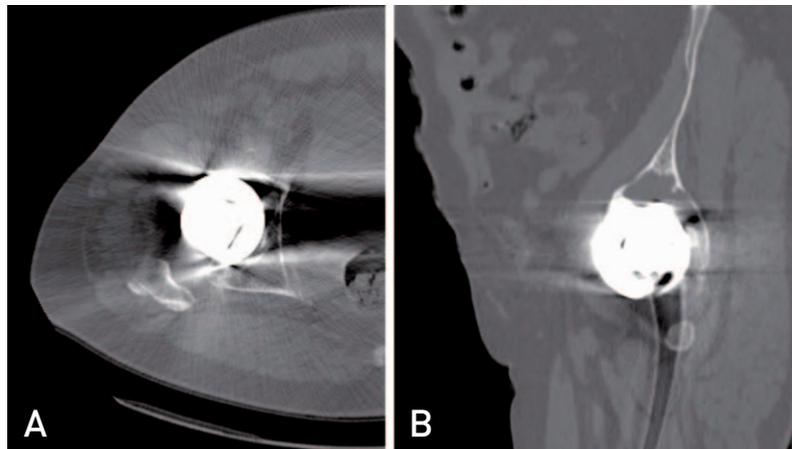
Fig. 2. Selected axial (**A**) and sagittal (**B**) computed tomography slices of a Paprosky type 2A acetabular defect.
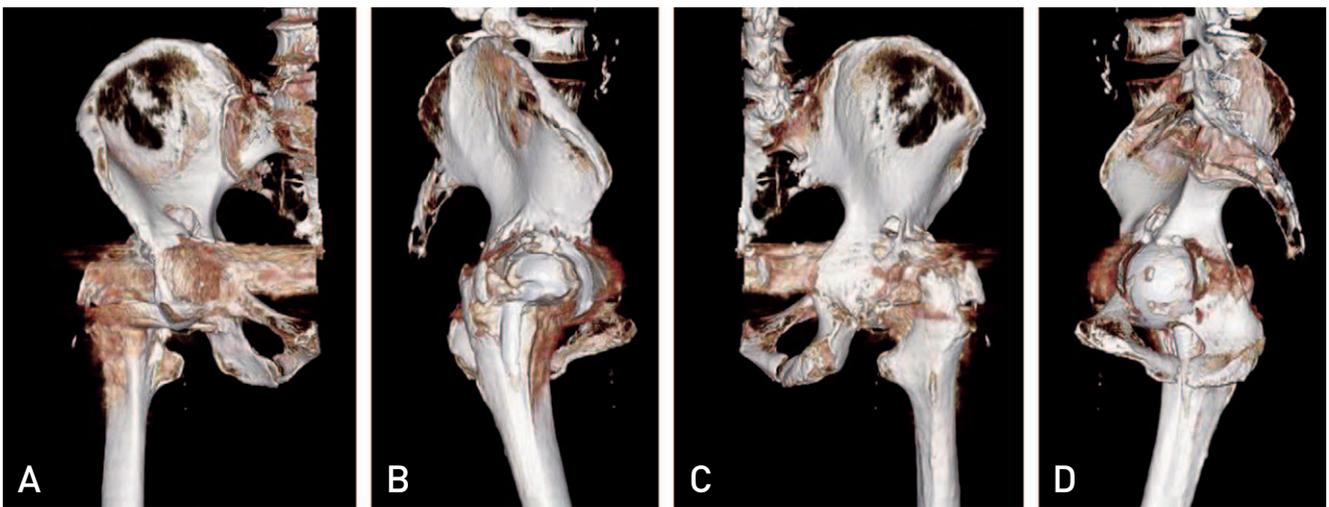
Fig. 3. Screenshots of selected frames (**A**: anterior, **B**: lateral, **C**: posterior, **D**: medial) from a three-dimensional reconstruction video of a Paprosky type 2A acetabular defect used in the presentation for this study.

modalities, level of training and defect type (SPSS software, ver. 19.0; IBM Co., Armonk, NY, USA)[9]. The approval from internal review board of Wake Forest Baptist Health was obtained (No. 18907).

RESULTS

Eight patients (5 males, 3 females) were identified that had plain radiographs, standard multiplanar CT scans, and 3D CT reconstructions available. The patient mean age was 76 years (range, 55-85 years) with a mean body mass index of 26.8 kg/m^2 (range, 23.3-33.7 kg/m^2). The correct classification regardless of imaging of PGY levels was 30% and overall interobserver reliability was poor (α=0.134, Table 2). Using X-ray led to 37% correctly identified defects, CT scans to 33% and 3D modeling to 20% of correct answers (P<0.001). Interobserver reliability was higher for X-rays (α=0.156) compared with CT scans (α=0.136) and 3D CT reconstruction (α=0.082), but there was poor interobserver reliability for all imaging modalities (Table 2). For type 1 defects, X-ray imaging had a significantly higher number of correct classification (92%) compared to CT scans (67%) and 3D modeling (31%, P<0.001, Table 3). Similarly, type 2A defects were classified correctly with

Table 2. Assessment of Interobserver Agreement Using Krippendorff' Alpha for Multiple Observers

Variable	Krippendorff's alpha*
Overall	0.134
Training level	
PGY-1	0.049
PGY-2	0.144
PGY-3	0.169
PGY-4	0.180
PGY-5	0.144
Fellows	0.181
Attendings	0.329
Imaging modality	
X-ray	0.156
2D CT	0.134
3D CT	0.082
Paprosky defect type	
Type 1	0.153
Type 2A	0.055
Type 2B	0.041
Type 2C	0.007
Type 3A	0.054
Type 3B	0.055

PGY: the postgraduate year of training, 2D: two-dimensional, 3D: three-dimensional, CT: computed tomography.
* 0-0.2: no agreement, 0.21-0.39: minimal, 0.40-0.59: weak, 0.60-0.79: moderate, 0.80-0.90: strong, >0.90: almost perfect.

Table 3. Correct Classification of Paprosky Type Defects Based on the Imaging Modality and in Relation to the Type of Defect

Variable	Classification						P-value
	Type I	Type IIA	Type IIB	Type IIC	Type IIIA	Type IIIB	
X-ray	92%	49%	39%	19%	17%	28%	<0.001
CT scan	67%	36%	31%	26%	24%	14%	<0.001
3D reconstructions	31%	15%	17%	31%	13%	8%	0.023
P-value	<0.001	0.007	0.108	0.245	0.450	0.056	

CT: computed tomography, 3D: three-dimensional.

Table 4. Correct Classification of Paprosky Type Defects Stratified by Level of Training and in Relation to the Type of Defect

Resident level	Classification						P-value
	Type I	Type IIA	Type IIB	Type IIC	Type IIIA	Type IIIB	
PGY-1	40%	53%	20%	24%	20%	20%	0.170
PGY-2	67%	40%	28%	20%	13%	13%	<0.001
PGY-3	64%	36%	31%	28%	21%	4%	<0.001
PGY-4	71%	24%	33%	26%	6%	19%	0.001
PGY-5	67%	20%	28%	28%	33%	13%	0.045
Fellows	83%	33%	50%	20%	33%	50%	0.227
Attendings	67%	0%	30%	10%	0%	33%	0.039
Overall	64%	33%	29%	25%	18%	16%	<0.001
P-value	0.518	0.216	0.757	0.894	0.316	0.161	

PGY: the postgraduate year of training.

Hip & Pelvis

Johannes F. Plate et al. Imaging Modalities to Identify Acetabular Defects in Residency Training

higher frequency on X-ray (49%) compared to CT scans (36%) or 3D reconstruction (15%, P=0.007). For type 2B, 2C, 3A and 3B defects, the type of imaging did not influence the frequency of correct answer. Interobserver reliability was poor for all Paprosky defect types (Table 2).

The level of training did not influence the frequency of correct classification regardless of the type of defect (P= 0.531). However, there was a significant difference in the frequency of correct classification based on the defect type (Table 4). With increasing severity of the bone defect, correct classification decreased. Regardless of level of training or imaging, 64% of observers recognized type 1 defects, compared to only 16% correct recognition of 3B defects. Interobserver reliability was poor for all levels of training and was minimal for attending surgeon (α=0.329, Table 2).

DISCUSSION

The current study revealed that classification of acetabular defects by orthopaedic residents using the Paprosky classification was unreliable. Overall, only 30% of acetabular defects were recognized correctly regardless of imaging modality and interobserver reliability was low. While the use of plain X-rays led to an increased number of correct classifications, using multiplanar CT scans or 3D CT reconstructions did not improve accuracy.

Previous studies have examined in the intra- and interobserver reliability of the Paprosky classification system. In a study by Campbell et al.[4] anteroposterior and Judet views of 33 hips in 30 patients with acetabular defects were assessed by 3 experienced orthopaedic surgeons, 3 senior orthopaedic residents, and 3 investigators that were part of the development of the classification system. All images were analyzed twice by each observer in a two-week period and graded using 3 different classification systems. There was moderate intraobserver agreement within the group of the classification developers for the Paprosky classification only, but poor intraobserver agreement within the orthopaedic experts and senior resident groups for all 3 classification system. Interobsever reliability was poor for all classification systems regardless of the level of training. In a study by Gozzard et al.[2], 25 anteroposterior and lateral X-rays of patients with acetabular bone loss were assessed by 4 observers using various classification systems including the Paprosky classification. The validity of the Paprosky classification was found to be good by comparing the preoperative grading and the intraoperative level of bone loss. However, there was only poor to moderate interobserver

reliability using the Paprosky classification. In a study by Yu et al.[10], 85 acetabular defects were reviewed by 4 observers including 3 consultant level orthopaedic surgeons and a medical graduate on 3 separate occasion. Two of the observers received teaching on the use of the classification system between the sessions. The study revealed that teaching can improve the ability to use the Paprosky classification system and can lead to good interobserver agreement (kappa= 0.65).

Further studies have sought to improve the classification of acetabular bone loss by including 3D image reconstructions or physical models based on preoperative CT scans. Munjal et al.[5] assessed the AAOS classification of acetabular defects using plain X-rays, CT scans, and CT-based 3D image reconstructions by one orthopaedic surgeon and one radiologist. Compared to intraoperative measurements, classification using plain X-rays and standard CT scans did not correlate, while using the 3D image reconstructions led to significant correlation[5]. Robertson et al.[11] compared the classification of acetabular bone loss on plain X-rays using the Paprosky classification to physical 3D models generated from preoperative CT scans by two independent reviewers. The interobserver reliability using plain X-rays was poor compared to good agreement using the physical 3D model.

The current study sought to improve acetabular defect classification by orthopaedic residents, fellows and attendings employing advanced imaging and 3D reconstructions. The use of 3D reconstructions did not improve correct classification but appeared to further decrease the ability to correctly identify defects. However, the study was limited by the relatively small number of participants and uneven representation of the level of training. While all observers received an educational lecture on the use of the classification system prior to reviewing the cases for the study, further teaching and repeat evaluation was not included in the study. The cases were chosen retrospectively and the "gold standard" was determined by two fellowship-trained adult reconstructive surgeons rather than intraoperative assessment. Cases were selected based on the availability of imaging modalities that most resembled each Paprosky type defect which could present selection bias. All Paprosky types and imaging modalities were available for each case to limit selection bias. While the number of cases is relatively small, different imaging modalities for similar defects were thereby used and therefore observers assessed the same pelvic defect using 3 different imaging modalities. The interobserver reliability assessment is limited by not taking the correctness of the answer into consideration. While there may be agreement by the observers, this could

relate to the wrong answer to be chosen more frequently than the correct answer. Observers were unable to scroll through the multiplanar CT scans which may have limited the ability to identify acetabular defects. However, this was similar to a test-taking environment at the time; the examination for the American Board of Orthopaedic Surgeons and orthopaedic surgery in-training examination did not provide scrollable CT scans. Furthermore, images were shown an arbitrary number of 60 seconds which was deemed as appropriate by the examiners and similar to an exam situation. Previous THA implants lead to substantial artefact during CT scans which decreases the quality of resulting 3D reconstructions and may impair visualization of crucial landmarks. This may be a general limitation of 3D reconstructions based on CT scans not inherent to the current study.

CONCLUSION

The findings of the current study suggest that further advanced imaging does not improve recognition of defects using the Paprosky classification. Therefore, the use of advanced imaging and associated cost increase and healthcare resource utilization should be questioned. With a predicted increase in the number of THA revisions in the future, increased resident education may be required during training to guide important surgical treatment decisions for these complex revision cases. The addition of 3D CT reconstructions did not appear to improve the residents' ability to identify defects correctly compared to plain X-rays.

CONFLICT OF INTEREST

The author declares that there is no potential conflict of interest relevant to this article.

REFERENCES

1. 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.*
2. 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;18:638-42.*
3. Parry MC, Whitehouse MR, Mehendale SA, et al. *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;20:50-5.*
4. Campbell DG, Garbuz DS, Masri BA, Duncan CP. *Reliability of acetabular bone defect classification systems in revision total hip arthroplasty. J Arthroplasty. 2001;16: 83-6.*
5. Munjal S, Leopold SS, Kornreich D, Shott S, Finn HA. *CT-generated 3-dimensional models for complex acetabular reconstruction. J Arthroplasty. 2000;15:644-53.*
6. Horas K, Arnholdt J, Steinert AF, Hoberg M, Rudert M, Holzapfel BM. *Acetabular defect classification in times of 3D imaging and patient-specific treatment protocols. Orthopade. 2017;46:168-78.*
7. Garrett J, Halvorson J, Carroll E, Webb LX. *Value of 3-D CT in classifying acetabular fractures during orthopedic residency training. Orthopedics. 2012;35:e615-20.*
8. Telleria JJ, Gee AO. *Classifications in brief: Paprosky classification of acetabular bone loss. Clin Orthop Relat Res. 2013;471:3725-30.*
9. Hayes AF, Krippendorff K. *Answering the call for a standard reliability measure for coding data. Commun Methods Meas 2007;1:77-89.*
10. Yu R, Hofstaetter JG, Sullivan T, Costi K, Howie DW, Solomon LB. *Validity and reliability of the Paprosky acetabular defect classification. Clin Orthop Relat Res. 2013;471:2259-65.*
11. Robertson DD, Sutherland CJ, Lopes T, Yuan J. *Preoperative description of severe acetabular defects caused by failed total hip replacement. J Comput Assist Tomogr. 1998;22: 444-9.*