

Arthroscopic Assessment of Cartilage Repair: A Validation Study of 2 Scoring Systems

Graham D. Smith, M.B., Ch.B., M.R.C.S., Jo Taylor, M.B., B.S., M.R.C.S.,
Karl Fredrik Almqvist, M.D., Ph.D., Christoph Erggelet, M.D., Ph.D., Gunnar Knutsen, M.D.,
Montse Garcia Portabella, M.D., Tony Smith, F.R.C.S., and James B. Richardson, M.D., F.R.C.S.

Purpose: This study tested the validity and reliability of the International Cartilage Repair Society (ICRS) cartilage repair assessment and the Oswestry Arthroscopy Score (OAS), which have been designed to assess repair of articular cartilage. **Type of Study:** Prospective validation study of arthroscopic cartilage repair scores. **Methods:** Arthroscopic videos were assessed by a panel of orthopaedic surgeons specializing in cartilage repair. Scoring was repeated after a 2-month interval. Scorers also answered a questionnaire to assess the face and content validity of the scoring systems. Validity of the 2 systems was compared and reliability and repeatability were measured. Pearson's correlation coefficient was used to measure equivalence reliability. The interclass correlation coefficient (ICC) was used to assess the repeatability and inter-rater reliability of each score, and internal consistency was assessed with Cronbach's alpha. **Results:** Face and content validity are acceptable for both scores. There is good agreement (equivalence reliability) between the scores (Pearson's correlation coefficient, $r = .88$; $P < .001$). Stability (interobserver reliability) and repeatability (test-retest reliability) are satisfactory for both scores with an ICC >0.7 for each score. Cronbach's alpha was 0.91 for ICRS and 0.82 for OAS, indicating better internal consistency for the ICRS score. **Conclusions:** The ICRS and OAS arthroscopic scores have been validated for the assessment of cartilage repair and both have been found to be statistically reliable and repeatable. The ICRS score does not allow for graft hypertrophy and may overscore in this situation, whereas the OAS includes assessment of graft stiffness. Both scores show satisfactory stability and repeatability. Internal consistency is adequate for both scores, although it is higher for the ICRS score. Both the ICRS and OAS arthroscopic scores are effective tools in the evaluation of cartilage repair. **Level of Evidence:** Level III, diagnostic study of nonconsecutive patients (no consistently applied reference gold standard). **Key Words:** Arthroscopy—Articular cartilage—Treatment outcome—Validation study—Reproducibility of results.

From the Robert Jones and Agnes Hunt Orthopaedic and District Hospital (G.D.S., J.T., J.B.R., T.S.), Oswestry, England; the Department of Physiotherapy and Orthopaedics, Ghent State University Hospital (K.F.A.), Ghent, Belgium; the Department of Orthopaedics, University of Freiberg (C.E.), Freiberg, Germany; the Department of Orthopaedics, University of Tromsø (G.K.), Tromsø, Norway; and the Department of Orthopaedics, Vall D'Hebron Hospital (M.G.P.), Barcelona, Spain.

Supported by a grant from EUROCELL (5th Framework EU QLGI-CT-2000-00496).

*Address correspondence and reprint requests to Graham D. Smith, M.B., Ch.B., M.R.C.S., Department of Trauma and Orthopaedics, Milton Keynes General Hospital, Standing Way, Eaglestone, Milton Keynes, Buckinghamshire, MK6 5LD, England.
E-mail: gdsmith@doctors.org.uk*

*© 2005 by the Arthroscopy Association of North America
0749-8063/05/2112-4231\$30.00/0
doi:10.1016/j.arthro.2005.09.007*

The emerging field of cartilage repair is a growing area in orthopaedics and is the subject of intense experimental and clinical research. Arthroscopic assessment of repair tissue in the defect is an important part of the evaluation of these treatments.

As part of the Euro-Cell collaboration on autologous chondrocyte implantation (ACI), the role of arthroscopy in the assessment of cartilage repair has been examined. Arthroscopic scoring systems to measure the quality of repair following articular cartilage surgery can be used as a semiquantitative outcome measure. In this study, the validity, reliability and repeatability of the International Cartilage Repair Society (ICRS) cartilage repair assessment and the Oswestry Arthroscopy Score (OAS) were compared.

Arthroscopy is the gold standard investigation of intra-articular pathology of the knee.¹ A number of

TABLE 1. *Oswestry Arthroscopy Score*

Graft level with surrounding cartilage	
Level	2
Raised	1
Below	0
Integration with surrounding cartilage	
Complete	2
Minor disruption (<25% of area)	1
Major disruption (>25% of area)	0
Appearance of surface	
Smooth	2
Fine fronds	1
Severe fronds/fibrillation	0
Color of graft	
Pearly, hyaline-like	2
White	1
Yellow bone	0
Stiffness on probing	
Normal compared to adjacent cartilage	2
Softer	1
Very soft/hard	0
Total	

scoring systems are in use that quantify or categorize the severity of cartilage damage.²⁻⁴ Surgical techniques aimed at repairing articular cartilage are becoming increasingly used, including ACI,⁵ microfracture,⁶ and mosaicplasty.⁷ Much work is being done to assess the effectiveness of these procedures. Previously available arthroscopic scoring systems focus on assessment of cartilage damage but are not readily applicable to measuring cartilage repair. Therefore, it has been necessary to devise new systems to assess cartilage repair that can be used as an outcome measure in the arthroscopic evaluation of cartilage repair techniques.

The ICRS system of assessment of cartilage repair has been developed from the work of Brittberg and Peterson, who were the first to perform ACI in humans.⁵ The Brittberg scoring system to evaluate macroscopic cartilage defect repair was described by Peterson et al. in 2000.⁸ This is a 3-point scoring system, each part being assigned a score between 0 and 4, giving an index with range 0 to 12 points (Table 2). This classification has been adopted by the International Cartilage Repair Society and used in the ICRS cartilage assessment.⁹ From this index, an overall grade can be deduced (normal, 12; nearly normal, 11-8; abnormal, 7-4; and severely abnormal, 3-0).

The OsCell group based in Oswestry, England, has been performing ACI since 1996 and has also developed an arthroscopic scoring system for cartilage repair, the OAS. This score includes the macroscopic

parameters that were felt to be most relevant to the quality of cartilage repair. It is divided into 5 parts, with each part awarded 0, 1, or 2 points. Points from each section are added to give an index with range 0 to 10 points (Table 1).

METHODS

Five arthroscopic video clips were circulated to 6 orthopaedic surgeons who have an interest in cartilage repair. The videos were made at the time of arthroscopic assessment in patients who had previously undergone ACI in the knee. The clips gave an overview of the whole of each treated area and the appearance on probing. The cases were selected to represent a spectrum of macroscopic appearance, from good to poor. Each surgeon was asked to score the videos using both the ICRS and OAS systems. After a period of 6 weeks, a second set of score forms was distributed to the panel of surgeons and the same videos were rescored. Although a power study was not performed, by using 5 videos and 6 surgeons, it was estimated that the total number of observations made (60) would be sufficient to provide meaningful results. A questionnaire was also circulated to the scorers to make a subjective assessment of the face and content validity of the 2 systems.

We performed statistical testing in terms of equivalence, stability, and interobserver reliability. We used

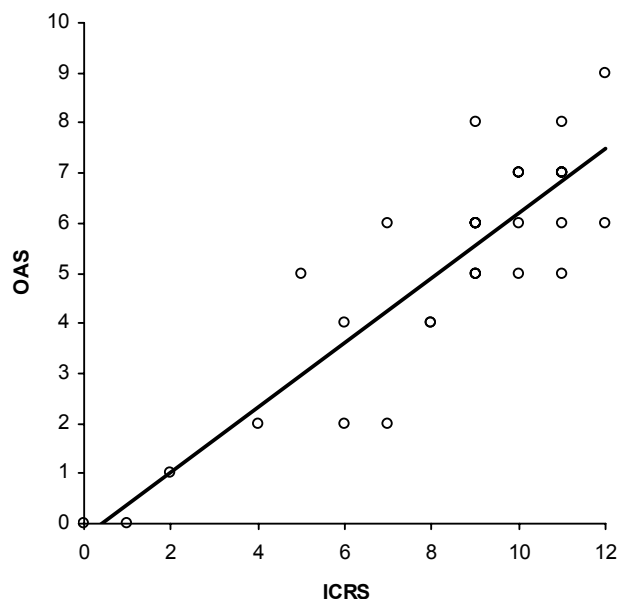


FIGURE 1. Scatter plot of ICRS cartilage assessment versus OAS. Pearson correlation coefficient = .88 (range, 0.76 to 0.94), $P < .01$.

TABLE 2. International Cartilage Repair Society (ICRS) Cartilage Repair Assessment

CARTILAGE REPAIR ASSESSMENT

Criteria	Points	
Degree of Defect Repair I Protocol A ⁽¹⁾	* In level with surrounding cartilage	4
	* 75% repair of defect depth	3
	* 50% repair of defect depth	2
	* 25% repair of defect depth	1
	* 0% repair of defect depth	0
I Protocol B ⁽²⁾	* 100% survival of initially grafted surface	4
	* 75% survival of initially grafted surface	3
	* 50% survival of initially grafted surface	2
	* 25% survival of initially grafted surface	1
	* 0% (plugs are lost or broken)	0
II Integration to Border zone	* Complete integration with surrounding cartilage	4
	* Demarcating border < 1mm	3
	* 3/4 of graft integrated, 1/4 with a notable border >1mm width	2
	* 1/2 of graft integrated with surrounding cartilage, 1/2 with a notable border > 1mm	1
	* From no contact to 1/4 of graft integrated with surrounding cartilage	0
III Macroscopic Appearance	* Intact smooth surface	4
	* Fibrillated surface	3
	* Small, scattered fissures or cracks	2
	* Several, small or few but large fissures	1
	* Total degeneration of grafted area	0
Overall Repair Assessment	Grade I normal	12 P
	Grade II nearly normal	11-8 P
	Grade III abnormal	7-4 P
	Grade IV severely abnormal	3-1 P

Cartilage Biopsy •

Location _____

(1) Protocol A:	(2) Protocol B:
autologous chondrocyte implantation (ACI); periosteal or perichondral transplantation; subchondral drilling; microfracturing; carbon fiber implants; others:	Mosaicplasty; OAT; osteochondral allografts; others:

TABLE 3. Validation Questionnaire for Arthroscopic Scores

I. Do you think that each method is a reasonable way to objectively assess cartilage repair arthroscopically?	
a. ICRS	Responses
i. = strongly disagree	0
ii. = disagree	0
iii. = undecided	1
iv. = agree	5
v. = strongly agree	0
b. Oswestry score	
i. = strongly disagree	0
ii. = disagree	0
iii. = undecided	0
iv. = agree	6
v. = strongly agree	0
II. Which parameters do you think are the most important in the assessment of cartilage repair? (Number in order of preference)	
	Median Rank
a. Thickness of repair	4
b. Color of repair	6
c. % area filled with repair tissue	2
d. Integration of repair tissue at border of lesion	2
e. Stiffness of repair tissue	3
f. Surface appearance	4.5

Pearson's correlation coefficient to measure equivalence. Stability and intraobserver reliability were assessed with the intraclass correlation coefficient (ICC). We used Cronbach's alpha as a measure of internal consistency.

RESULTS

Validity

To test the validity of the 2 scores, the panel of scorers was given a questionnaire regarding the scoring systems and the content of these scores. The questionnaire and responses are summarized in Table 3.

Face Validity: Face validity assesses whether a test appears to be a reasonable way of gaining the information required. From the questionnaire, all but 1 scorer answered that they "agreed" that the ICRS was a reasonable method of assessing articular cartilage repair; the other scorer was undecided. All scorers "agreed" that the OAS was a reasonable method.

Content Validity: Content validity looks at whether a scoring system is representative of the parameter being measured. Both ICRS and OAS scores measure the degree of repair, integration of the edges of the graft, and the appearance of surface. The OAS

also measures color and stiffness of the graft. When scorers were asked to rank which parameters were the most important in the arthroscopic assessment of cartilage repair, percentage area filled and border integration were ranked the equal most important factor, followed by stiffness, then thickness, and then surface appearance. Color was unanimously agreed to be the least important factor. Therefore, of the 2 additional parameters used in the OAS, only the stiffness was thought by the reviewers to be of importance in assessing cartilage repair. The ICRS score gives no allowance for hypertrophy of a grafted area, which is a recognized complication of ACI and may be symptomatic. In this situation, defect filling would get full marks on the degree of repair section of the ICRS score but would lose a mark on the OAS score. Therefore, one could expect a hypertrophied graft to score better on the ICRS score.

Reliability

Equivalence Reliability: The scores from each patient were plotted against each other (Fig 1). Pearson's correlation coefficient $r = .88$ (range, .76 to .94), $P < .01$, indicated good equivalence of the 2 scores.

Stability and Inter-rater Reliability: The ICC is used to measure the stability and inter-rater reliability. ICC is 0.94 (ICRS) and 0.94 (OAS) for test-retest reliability and 0.83 (ICRS) and 0.76 (OAS) for inter-rater reliability. A level of >0.7 is taken to represent adequate reliability.¹⁰ Both the ICRS and OsCell score meet this criterion for both stability (test-retest) and inter-rater reliability.

Internal Consistency

Internal consistency looks at whether all items on a scale are correlated, so they measure the same thing. This can be measured using by using Cronbach's alpha.¹¹ A value of α between 0.7 and 0.8 is satisfactory for comparing groups as a research tool, although α should be greater than 0.9 for use as a clinical tool for individuals.¹² Cronbach's alpha is calculated as 0.82 for OAS and 0.91 for ICRS. Therefore, both scores have satisfactory internal consistency for use in research, but the ICRS score, with $\alpha = 0.91$, might be better for measurement in individual patients.

DISCUSSION

This study has shown that both the ICRS and OAS are statistically reliable and valid for the purposes of assessing articular cartilage repair following ACI. There are advantages and disadvantages with each

system. This study has shown that both tests have statistical equivalence and that they both have reasonable stability and inter-rater reliability. Better internal consistency characteristics might suggest that the ICRS score is suited for individual patient assessment, although Cronbach's alpha is satisfactory for both scores to be used to measure groups of patients in a clinical research setting. The OAS makes an allowance for hypertrophy of the graft. This factor is not accounted for in the ICRS assessment and may be of clinical importance. The ICRS system is in more common use and a number of clinical studies have used this system.^{8,13,14}

Assessment of arthroscopic video clips could be criticized as being less accurate in assessing cartilage repair than arthroscopic assessment by an individual surgeon. To carry out this study, the only feasible way to get a number of surgeons to assess the same joint so as to measure interobserver reliability was to use video footage of previously performed arthroscopies. The findings of good interobserver and intraobserver reliability suggest that this is a reasonable method.

This study was relatively small in scale, with only 5 videos being assessed. A power study was not performed and this number was chosen pragmatically so that participation by the 6 scorers was not unduly time consuming. A total of 60 observations were made in this study, and the values of the ICC give some reasonable values. This study does not discern whether 1 score is significantly superior to the other, but can show that they are both statistically reliable.

The available evidence from randomized clinical trials comparing different cartilage repair techniques is contradictory.¹⁴⁻¹⁶ These studies have reported short-term clinical results and no comparative data of longer term follow-up is available at present. Techniques such as ACI, microfracture, and mosaicplasty must not only be shown to be successful treatments in the short term, but also that they have long-lasting benefits. Measurement of the quality of cartilage repair in these patients is needed in addition to clinical assessment. Noninvasive investigations using magnetic resonance imaging have also been performed,^{13,17} although the presence of changes on magnetic resonance imaging, such as "marrow edema," is not fully understood. Performing arthroscopy as a follow-up procedure gives information on the macroscopic appearance of repair and may also allow biopsy specimens to be taken for histologic analysis. Histology remains the gold standard, and the aim of any technique is to achieve cartilage repair that is structurally and histochemically as near to normal hyaline

cartilage as possible. A biopsy specimen, which is usually taken from the center of an area of cartilage repair, only gives an impression from a small part of an area of cartilage repair as a whole. Therefore, it is important to use arthroscopic scoring systems in conjunction with histologic examination to get an overall impression of the quality of cartilage repair. Macroscopic assessment, including evaluating integration at the interface of the graft with surrounding normal cartilage, is just as important as analysis of the graft's microscopic structure. By using arthroscopic scoring with other qualitative measures, it may be possible to predict which treatments will give the best long-term outcome.

Acknowledgment: The authors thank Professor Peter Jones, Department of Statistics, Keele University, Staffordshire, England, for help with statistical methods and analysis.

REFERENCES

1. DeHaven KE, Collins HR. Diagnosis of internal derangements of the knee. The role of arthroscopy. *J Bone Joint Surg Am* 1975;57:802-810.
2. Outerbridge RE. The etiology of chondromalacia patellae. *Clin Orthop* 2001;389:5-8.
3. Dougados M, Ayrat X, Listrat V, et al. The SFA system for assessing articular cartilage lesions at arthroscopy of the knee. *Arthroscopy* 1994;10:69-77.
4. Noyes FR, Stabler CL. A system for grading articular cartilage lesions at arthroscopy. *Am J Sports Med* 1989;17:505-513.
5. Brittberg M, Lindahl A, Nilsson A, Ohlsson C, Isaksson O, Peterson L. Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. *N Engl J Med* 1994;331:889-895.
6. Steadman JR, Rodkey WG, Singleton SB, Briggs KK. Microfracture technique for full-thickness chondral defects: Technique and clinical results. *Oper Tech Orthop* 1997;7:300-304.
7. Hangody L, Kish G, Karpati Z. Arthroscopic autogenous osteochondral mosaicplasty—A multicentric, comparative, prospective study. *Index Traumatol Sport* 1998;5:3-9.
8. Peterson L, Minas T, Brittberg M, Nilsson A, Sjogren-Jansson E, Lindahl A. Two- to 9-year outcome after autologous chondrocyte transplantation of the knee. *Clin Orthop Rel Res* 2000;374:212-234.
9. Brittberg M, Winalski CS. Evaluation of cartilage injuries and repair. *J Bone Joint Surg Am* 2003;85:58-69.
10. Deyo RA, Diehri P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;4 Suppl:142S-158S.
11. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-333.
12. Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997;314:572.
13. Henderson IJ, Tuy B, Connell D, Oakes B, Hettwer WH. Prospective clinical study of autologous chondrocyte implantation and correlation with MRI at three and 12 months. *J Bone Joint Surg Br* 2003;85:1060-1066.
14. Bentley G, Biant LC, Carrington RW, et al. A prospective, randomised comparison of autologous chondrocyte implantation versus mosaicplasty for osteochondral defects in the knee. *J Bone Joint Surg Br* 2003;85:223-230.

15. Horas U, Pelinkovic D, Herr G, Aigner T, Schnettler R. Autologous chondrocyte implantation and osteochondral cylinder transplantation in cartilage repair of the knee joint. A prospective, comparative trial. *J Bone Joint Surg Am* 2003;85:185-192.
16. Knutsen G, Engebretsen L, Ludvigsen T, Drogset J, Grontvedt T, Solheim E, Strand T, et al. Autologous chondrocyte implantation compared with microfracture in the knee. A randomized trial. Presented at the International Cartilage Repair Society 4th Symposium, Toronto, Ontario, Canada, June 2002.
17. Burstein D, Gray M. New MRI techniques for imaging Cartilage. *J Bone Joint Surg Am* 2003;85:70-77.