Middleton, R. M., Vo, A., Ferguson, J., Judge, A., Alvand, A., Price, A. J., & Rees, J. L. (2017). Can Surgical Trainees Achieve Arthroscopic Competence at the End of Training Programs? A Cross-sectional Study Highlighting the Impact of Working Time Directives. *Arthroscopy*, *33*(6), 1151-1158. https://doi.org/10.1016/j.arthro.2016.10.025

Can surgical trainees achieve arthroscopic competence at the end of training programs: a cross sectional study highlighting the impact of working time directives.

ABSTRACT

Purpose

Our objective was to provide training guidance on procedure numbers by assessing how the number of previously performed arthroscopic procedures related to both competent and expert performance in simulated arthroscopic shoulder tasks. The null hypothesis was that the recommended minimum number of arthroscopic cases in UK training are adequate to achieve competency.

Methods

A cross-sectional study assessing simulated shoulder arthroscopic performance was undertaken. 45 participants of varying experience performed two validated tasks: a simple diagnostic task and a more complex Bankart labral repair task. All participants provided logbook numbers for previously performed arthroscopies. Performance was assessed using a Global Rating Scale (GRS) and motion analysis. Receiver operating characteristic (ROC) curve analyses were conducted to identify optimum cut points for task proficiency at both 'competent' and 'expert' levels.

Results

Increasing surgical experience resulted in significantly better performance for both tasks as assessed by GRS or motion analysis ($p<0.0001$). ROC curve analyses demonstrated 52 previous arthroscopies were needed to perform to a 'competent' level at the diagnostic task and 248 to be 'competent' at the complex task. To perform at an expert level, 290 and 476 previous arthroscopies respectively were needed.

Conclusions

This study provides quantified guidance for arthroscopic training and highlights the positive relationship between arthroscopic case load and arthroscopic competency. We have estimated that the number of arthroscopies required to achieve competency in a basic arthroscopic task exceed those recommended in some countries. . These estimates provide useful guidance to those responsible for training programmes.

Clinical Relevance

The numbers to achieve competent arthroscopic performance in the assessed simulated tasks exceed what is recommended and what is possible during surgical training programmes in some countries.

INTRODUCTION

Traditionally orthopaedic surgical training has been delivered through an apprenticeship model with trainees increasing their level of involvement in surgical procedures at the discretion of their trainers.[1,2] The introduction of working time directives have resulted in a relatively unpublicised yet dramatic reduction in training time in some countries, including the United Kingdom (UK)[3–6]. Similar restrictions introduced in the United States of America by the Accreditation Council for Graduate Medical Education (ACGME) in 2011 have raised concerns regarding patient outcomes and resident education, with little overall difference observed in resident well-being.[7–9]

The introduction of minimally invasive surgery in some specialties has resulted in higher complication rates during initial skill acquisition, with some trainees finding these new skills difficult to acquire.[10–14] In orthopaedics an increasing number of procedures are performed arthroscopically. It is acknowledged that while a substantial number of repetitions are required to attain arthroscopic competency, the exact numbers are unknown. It is unlikely there is one ‘fixed number’ because of the variety of arthroscopic operations and variations in the innate skills of trainees.[15] In the UK trainee surgeons must perform a minimum of 40 arthroscopic procedures during their 6 year training programme. This target makes no distinction between joint or complexity, and is not clearly based on any particular evidence.[16] Our group has recently published estimates for the number of diagnostic knee arthroscopies required to reach competent and expert proficiency.[17] However we did not explore this issue with regards to a more complex task. Such guidance would be of value to those setting national standards, curricula and running residency training programmes in the face of working time restrictions.

Our objective was to provide training guidance on procedure numbers by assessing how the number of previously performed arthroscopic procedures related to both competent and expert performance in simulated arthroscopic shoulder tasks. The null hypothesis was that the recommended minimum number of arthroscopic cases in UK training are adequate to achieve competency.

METHODS

A cross-sectional study was undertaken to investigate the impact of previous arthroscopic case load on the ability to achieve competency on a simple and a complex simulated arthroscopic shoulder task.

Participants

Institutional review board approval was granted for this non-patient study. Each candidate was provided with an information sheet and gave informed consent. Over a four month period, 45 participants with varying degrees of arthroscopic experience were invited to take part in this cross sectional training study. In addition to faculty members, current medical students, interns and orthopaedic trainees rotating though our institution were eligible for inclusion. Exclusion criteria included prior exposure to the arthroscopic task and simulator used in this study. There were 14 novices with no arthroscopic experience (medical students and interns), 27 orthopaedic trainees of varying experience and 4 expert orthopaedic surgeons with a specialist shoulder interest were recruited on statistical advice so to enable the planned ROC analysis. The numbers of all previous arthroscopies performed were collected via surgical logbooks for each participant. All previous arthroscopies performed were included but unsurprisingly for the trainee group, these were made up mainly of knee arthroscopies and to a lesser extent shoulder arthroscopies.

Simulator-based assessment of surgical skill

The simulated arthroscopic tasks were conducted in a surgical skills laboratory using the Alex Shoulder Professor II benchtop model (Sawbones Europe, Malmö, Sweden). This is a life-sized simulator with arthroscopic portals that allow conventional arthroscopic instruments to be used in a beech chair position. A standard 30 degree arthroscope, camera and high definition display were used for all cases (Smith and Nephew Endoscopy, Huntingdon, UK).

The study involved two previously validated simulated tasks (20,27). The first task was a more straight forward diagnostic triangulation task involving a systematic exploration of the shoulder using the arthroscopic probe to touch a series of marked points within the model. These points were numbered from 1-9 and corresponded to important landmarks visualised during real-life diagnostic arthroscopies of the shoulder (see figure 1).

The second task was a simulated arthroscopic Bankart (labral) repair. This involved fixing a detached labrum to the glenoid by screwing a single loaded suture anchor into the glenoid, passing a suture arthroscopically through the labrum with a suture passer and then tying a secure arthroscopic knot using a knot pusher and finally using a suture cutter to divide the suture ends. The model was prepared with a standardized predrilled hole in the peripheral anterior glenoid to accept the suture anchor.

Each task was preceded by a video presentation outlining the technique. It was watched once by each of the participants. Following the video there was a short opportunity for participants to practice tying their knots on a separate bench top model. All the tasks had standardised starting and finishing positions.

Motion Analysis

A three-dimensional electromagnetic motion tracking system (PATRIOT; Polhemus, Colchester, Vermont) was used to objectively monitor surgical performance. This system uses a fixed emitter and two small electromagnetic sensors worn on the dorsum of both hands. The movement information is automatically converted using custom software (MATLAB, version 6.5; The MathWorks, Natick, Massachusetts) to provide three outcome measures; 'total path length' (distance moved by the hands during the task); 'number of hand movements' performed and the overall 'time taken' to complete the task. This objective technique has been extensively validated for skill assessment in arthroscopy[18–22].

Global Rating Scale (GRS)

A validated [22,23] GRS was used to rate the performance of each candidate for the two different tasks. It uses six testable domains measuring instrument handling, depth perception, bimanual dexterity, flow of operation, efficiency, and quality of the final product. Each domain was rated using a Likert scale with anchors at 1, 3 and 5 points, giving a maximum possible score of 30 points and a minimum of 6. An independent observer not involved in collecting the data assigned the GRS based on the blinded video recordings taken automatically by the arthroscopic equipment and an external web cam. These video files were anonymised via a random number allocation and the webcam videos cropped so that any identifiable features were obscured. A proportion of these videos were rated by a further blinded assessor for inter-observer variability. For the Bankart task, the adequacy of the repair was also assessed using a pass/fail assessment, with a fail being awarded if there was cut-through of the suture, incorrect positioning of the anchor, a loose knot, or if there was gapping of the repair on probing of the repair after the task.

The primary outcome measure was assessment of arthroscopic performance using a validated GRS. This was then used to estimate the number of arthroscopies required to achieve competent and expert proficiency in the simulated tasks.

Statistical Analysis

Non-parametric tests were used as the data was not normally distributed according to the Kolmogorov-Smirnov test. Inter-rater reliability for GRS scoring was determined using the Cronbach alpha coefficient. The Kruskal-Wallis test was used to determine the differences between groups for each simulation task. Where a significant difference was found, pairwise comparisons between the groups were made using the Mann–Whitney U test. The Spearman rank correlation coefficient was used to analyse the correlation between the GRS and motion analysis parameters as well as performance between the two different simulation tasks.

To assess the importance and influence of number of previous procedures performed on GRS, ROC analyses were performed for each arthroscopic task. A score of 5/5 in each domain (30/30) was set as an ‘expert’ performance. This essentially reflected a flawless arthroscopic performance in the simulated setting. A score of 4/5 in each domain (total 24/30) was set as a ‘competent’ performance threshold as this was still judged to reflect a safe arthroscopic performance at these particular tasks with minimal errors and perhaps more representative of the performance levels actually achieved by most by the end of residency programmes. The optimum point on the ROC curve for sensitivity and specificity was determined, and number of cases at this point recorded. A p-value of 0.05 was considered statistically significant.

Source of Funding

This project was supported by the Association for Simulated Practice in Healthcare (ASPiH), a CAE Healthcare grant and by the National Institute of Health Research Oxford Biomedical Research Unit.

RESULTS

45 individuals participated in this study (14 novices, 27 trainees and 4 experts). No participants were excluded or failed to complete the study. While all 45 participants were able to perform the simple diagnostic task. 9 of the novices wereunable to complete the Bankart task due to it's complexityand so worst case scores were allocated to these individuals. There was an expected broad range of previous arthroscopic experience in the trainee group (table 1), with the vast majority of prior arthroscopic experience gained outside of the shoulder, predominantly knee arthroscopy.

Global Rating Scale

Two observers individually assessed a random selection of 16 videos (19.8% of the total) and excellent inter-rater reliability was noted (Cronbach $\alpha = 0.86$). Thereafter a single rater went on to score the GRS for the remaining 65 videos. The Kruskall-Wallis testdemonstrated significant differences in GRS scores for both the diagnostic triangulation task ($p = 0.0001$) and the Bankart task ($p = 0.0001$) (figure 2). Pair-wise comparisons of GRS performances between experience groups using the Mann-Whitney U test demonstrated significant differences between each group (table 2).

Motion Analysis.

All motion parameters improved for both tasks, across groups with increasing experience. The Kruskall-Wallis test demonstrated significant differences (p = 0.0001) for all metrics across all groups. Figure 3 is a graphical representation of the difference in performance between the 3 groups in terms of 'hand movement' motion analysis metrics. A similar pattern was observed for 'time taken' and 'total path length'. Pair-wise comparisons of motion data using the Mann-Whitney U test demonstrated significant differences between each group for each task (table 2), further demonstrating and confirming the construct validity of these tasks.

Impact of Previous Arthroscopic Experience

The number of previous arthroscopies performed by every participant was strongly correlated with performance on the arthroscopic task using GRS (table 3).

Number of arthroscopic procedures required in order to reach different levels of performance

Receiver operating characteristic (ROC) analyses were performed to determine 'minimum numbers' of previous arthroscopic procedures required to reach a competent level and expert level as previously defined in the statistical analysis section. ROC curve analyses demonstrated 52 previous arthroscopies were needed to perform to a 'competent' level at the diagnostic task and 248 to be 'competent' at the complex task. To perform at an expert level, 290 and 476 previous arthroscopies respectively were needed (table 4).

DISCUSSION

This study supports the accepted correlation between previous arthroscopic experience and performance on a given arthroscopic task. Using ROC analysis we have estimated the number of arthroscopic operations trainees need to perform in order to reach a given level of performance in two simulated tasks. Of concern, the number to achieve competency in a simple shoulder task exceeds the current minimum number in some countries. In recent years there has been increased emphasis on competency and quality of training, and less on operative numbers. Given the relationship between case load and performance, a role for minimum recommended cases performed during training remains relevant. ..Our aim was not to provide absolute figures, but rather provide guidance to the surgical education community. Although the ROC curve sensitivity and specificities for the 'competent' level performance estimates were not as high as the expert level estimates, the Area Under the Curve (AUC) for all analyses were greater than 0.8 implying reliable results, especially if they are taken as a guide. The results highlight the importance of task complexity on both competent and expert performance level. The finding that more complex tasks require greater experience will not be of surprise to most. Most educators would also accept that whilst achieving expert performance is the ideal outcome, surgical training programmes are usually only able to produce 'safe and competent' surgeons, with expert skills continuing to develop post training.

In the UK, i orthopaedic trainees are required to perform a minimum of 40 'arthroscopic operations' by the end of training. It is unclear how these numbers have been decided, but the suggestion is that they are deemed sufficient for trainees to progress to independent practice. Our study estimates 52 arthroscopies are needed just to perform the more simple diagnostic shoulder arthroscopy to a 'competent' level. These results therefore cast doubt on whether a minimum number of 40 total arthroscopies is adequate for independent practice. Furthermore, attaining competence in basic arthroscopy does not imply seamless progression to more complex therapeutic tasks, which often have significant learning curves. For example, the estimated cases required to attain a competent level of performance in the more complex Bankart task is even less achievable.

This study highlights the impact of recent changes to surgical training. Achieving the number of cases requiredto reach competent levels of performance are no longer possible, in part due to the reduction in surgical training time in some countries.[3,4,5] This was highlighted by a recent national interrogation of current UK surgical trainee logbooks. The maximum mean number of arthroscopic procedures by training region was recorded at 82, with the average number of procedures being half this, and the minimum being 28.[24] Given these numbers, our study suggests it is extremely difficult to attain 'competent' performance levels within the current training framework. As working time regulations seem unlikely to be eased, alternative solutions need to be found. Logically this shortfall can only be addressed by additional training such as additional fellowships post training, or reinforcing arthroscopic learning and training outside of the operating theatre using valid simulation methods.

Simulation based training seems likely to be a vital new adjunct in surgical training. Prior work has demonstrated improved learning and performance through simulation based training[21,22,25–27] There is now a growing argument that simulation should be integrated into training programmes to improve learning, and to help reduce any skills gap. Within the UK, simulation is set to become part of the orthopaedic curriculum, although no realistic resources are yet in place to support such a move.. A similar recognition of the importance of simulation in training has been highlighted in the USA[28], with collaboration between the American Academy of Orthopaedic Surgeons, the Arthroscopy Association of North America and the American Board of Orthopaedic Surgery to evaluate and improve the delivery of arthroscopic training using simulation. The Fundamentals in Arthroscopic Surgery Training (FAST) program is one example which aims to develop a recognised arthroscopic education programme for generic arthroscopic skills.[29] The Copernicus Initiative has also demonstrated the effectiveness of proficiency based training with simulation programmes for Bankart repair.[30] The experience and success of such projects may potentially encourage their expansion.

LIMITATIONS

The present study focused on shoulder arthroscopy, and provides guidance for two specific tasks, and may be criticised as being joint specific. Other arthroscopic procedures and joints present their own challenges, but it is likely that similarly high numbers of procedures are required to achieve competent and expert performance. Our study assessed performance on arthroscopic shoulder tasks, but appreciate that the majority of arthroscopic experience for trainees was gained from knee arthroscopy. However, we feel a number of generic arthroscopic skills are common to both joints. Therefore significant experience in knee arthroscopy is likely to translate to improved performance in shoulder arthroscopy. A final limitation is the accuracy of trainees' surgical logbooks. The majority of logbooks do not document the proficiency with which a case was performed. They may also contain inaccurate records – both as to the proportion of a procedure performed, and the number of cases performed.

CONCLUSIONS

This study provides quantified guidance for arthroscopic training and highlights the positive relationship between arthroscopic case load and arthroscopic competency. We have estimated that the number of arthroscopies required to achieve competency in a basic arthroscopic task exceed those recommended in some countries. . These estimates provide useful guidance to those responsible for training programmes.

References

1. Pedowitz RA, Esch J, Snyder S. Evaluation of a virtual reality simulator for arthroscopy skills development. *Arthrosc - J Arthrosc Relat Surg*. 2002;18(6). doi:10.1053/jars.2002.33791.
2. Michelson JD. Simulation in orthopaedic education: an overview of theory and practice. *J Bone Joint Surg Am*. 2006;88(6):1405–1411. doi:10.2106/JBJS.F.00027.
3. Benes V. The European Working Time Directive and the effects on training of surgical specialists (doctors in training): A position paper of the surgical diciplines of the countries of the EU. *Acta Neurochir (Wien)*. 2006;148(12):1326–1330. doi:10.1007/s00701-006-0880-8.
4. Giles JA. Surgical training and the European Working Time Directive: The role of informal workplace learning. *Int J Surg*. 2010;8(3):179–180. doi:10.1016/j.ijsu.2010.01.011.
5. Chikwe J, De Souza A, Pepper J. No time to train the surgeons. *Br Med J*. 2004;328(7437):418–9.
6. Greenaway D. *Shape of Training Review - Securing the future of excellent patient care*.; 2013.
7. Scally CP, Ryan AM, Thumma JR, Gauger PG, Dimick JB. Early impact of the 2011 ACGME duty hour regulations on surgical outcomes. *Surgery*. 2015:1–9. doi:10.1016/j.surg.2015.05.002.
8. Babu R, Thomas S, Hazzard MA, et al. Morbidity, mortality, and health care costs for patients undergoing spine surgery following the ACGME resident duty-hour reform:

Clinical article. *J Neurosurg Spine*. 2014;21(4):502–15. doi:10.3171/2014.5.SPINE13283.

9. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg*. 2014;259(6):1041–53. doi:10.1097/SLA.0000000000000595.

10. Allum R. Complications of Arthroscopy of the Knee. *J Bone Jt Surg (British Vol*. 2002;84(September):937–945.

11. Hanna GB, Shimi SM, Cuschieri A. Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy. *Lancet (London, England)*. 1998;351(9098):248–51. doi:10.1016/S0140-6736(97)08005-7.

12. Cuschieri A. Whither minimal access surgery: tribulations and expectations. *Am J Surg*. 1995;169(1):9–19. doi:10.1016/S0002-9610(99)80104-4.

13. Deziel DJ, Millikan KW, Economou SG, Doolas A, Ko ST, Airan MC. Complications of laparoscopic cholecystectomy: a national survey of 4,292 hospitals and an analysis of 77,604 cases. *Am J Surg*. 1993;165(1):9–14. doi:10.1016/S0002-9610(05)80397-6.

14. Hunter JG, Sackier JM, Berci G. Training in laparoscopic cholecystectomy - Quantifying the learning curve. *Surg Endosc*. 1994;8(1):28–31. doi:10.1007/BF02909489.

15. O’Neill PJ, Cosgarea AJ, Freedman JA, Queale WS, McFarland EG. Arthroscopic proficiency: A survey of orthopaedic sports medicine fellowship directors and orthopaedic surgery department chairs. *Arthroscopy*. 2002;18(7):795–800. doi:10.1053/jars.2002.31699.

16. Joint Committee on Surgical Training. Guidelines for the award of a CCT in Trauma & Orthopaedic Surgery. 2015. Available at: http://www.jcst.org/quality-

assurance/documents/certification-guidelines/trauma-and-orthopaedic-surgery-certification-guidelines.

17. Price AJ, Erturan G, Akhtar K, Judge A, Alvand A, Rees JL. Evidence-based surgical training in orthopaedics: how many arthroscopies of the knee are needed to achieve consultant level performance? *Bone Joint J*. 2015;97-B(10):1309–15. doi:10.1302/0301-620X.97B10.35973.

18. Howells NR, Brinsden MD, Gill RS, Carr AJ, Rees JL. Motion Analysis: A Validated Method for Showing Skill Levels in Arthroscopy. *Arthrosc - J Arthrosc Relat Surg*. 2008;24(3):335–342. doi:10.1016/j.arthro.2007.08.033.

19. Alvand A, Auplish S, Khan T, Gill HS, Rees JL. Identifying orthopaedic surgeons of the future: the inability of some medical students to achieve competence in basic arthroscopic tasks despite training: a randomised study. *J Bone Joint Surg Br*. 2011;93(12):1586–91. doi:10.1302/0301-620X.93B12.27946.

20. Pollard TC, Khan T, Price AJ, Gill HS, Glyn-Jones S, Rees JL. Simulated hip arthroscopy skills: learning curves with the lateral and supine patient positions : a randomized trial. *J Bone Jt Surgery Am Vol*. 2012;94(10):e68 1–10. doi:10.2106/JBJS.K.00690.

21. Jackson WFM, Khan T, Alvand A, et al. Learning and retaining simulated arthroscopic meniscal repair skills. *J Bone Jt Surgery Am Vol*. 2012;94(17):e132:1–8. doi:10.2106/JBJS.K.01438.

22. Alvand A, Logishetty K, Middleton R, et al. Validating a global rating scale to monitor individual resident learning curves during arthroscopic knee meniscal repair. *Arthrosc - J Arthrosc Relat Surg*. 2013;29(5):906–912. doi:10.1016/j.arthro.2013.01.026.

23. Insel A, Carofino B, Leger R, Arciero R, Mazzocca AD. The development of an objective

model to assess arthroscopic performance. *J Bone Joint Surg Am*. 2009;91(9):2287–2295. doi:10.2106/JBJS.H.01762.

24. Penn-Barwell J, Bennett P, Wood A, Reed M. Is there variation in orthopaedic operative training experience between deaneries? *Bone Jt J*. 2015;97-B(8):22.

25. Howells NR, Auplish S, Hand GC, Gill HS, Carr AJ, Rees JL. Retention of arthroscopic shoulder skills learned with use of a simulator. Demonstration of a learning curve and loss of performance level after a time delay. *J Bone Joint Surg Am*. 2009;91(5):1207–1213. doi:10.2106/JBJS.H.00509.

26. Slade Shantz JA, Leiter JRS, Gottschalk T, MacDonald PB. The internal validity of arthroscopic simulators and their effectiveness in arthroscopic education. *Knee Surgery, Sport Traumatol Arthrosc*. 2014;22(1):33–40. doi:10.1007/s00167-012-2228-7.

27. Howells NR, Gill HS, Carr AJ, Price AJ, Rees JL. Transferring simulated arthroscopic skills to the operating theatre: a randomised blinded study. *J Bone Joint Surg Br*. 2008;90(4):494–499. doi:10.1302/0301-620X.90B4.20414.

28. Lubowitz JH, Provencher MT, Brand JC, Rossi MJ. You…Can Be a Millionaire. *Arthrosc J Arthrosc Relat Surg*. 2015;31(2):177–180. doi:10.1016/j.arthro.2014.12.001.

29. Surgical Skills Task Force of the American Board of Orthopaedic Surgery. ABOS Module #14 Basic Arthroscopy Skills. Available at: https://www.abos.org/media/7061/module_14_basic_arthroscopy_skills_edit.pdf. Accessed August 15, 2015.

30. Angelo RL, Ryu RKN, Pedowitz RA, Gallagher A. Results from the Arthroscopic Association of North America (AANA) Copernicus Initiative; A Multicentre Prospective, Randomized, Blinded Trial of Proficiency-Based Progression Training Employing

Simulation for an Arthroscopic Bankhart Procedure. In: *ISAKOS Congress*. Lyon, France; 2015.

Figure Legend

Figure 1. Arthroscopic view for diagnostic task showing numbered point and arthroscopic probe (left glenohumeral joint).

Figure 2. Box and whisker plots demonstrating performance in the two tasks as assessed by GRS scores for each of the subject groups (Novice, Trainee and Expert). (Note: In all plots, the box represents the interquartile range, the bold line in the box represents the median, and the whiskers represent the 2.5 and 97.5 percentiles. Outlier values are represented by the small circles)

Figure 3: Box and whisker plots demonstrating the difference in hand movements between the groups (Novice, Trainee and expert) for the two tasks. (Note: In all plots, the box represents the interquartile range, the bold line in the box represents the median, and the whiskers represent the 2.5 and 97.5 percentiles. The outlier values are represented by the small circles)

Tables

	Total arthroscopies	Shoulder arthroscopies	Non-shoulder arthroscopies
Novices	0	0	0
Trainees	302.9 (18.0 – 975.0)	42.2 (0.0 - 347.0)	260.4 (18.0 - 965.0)
Experts	1429.8 (830 – 2035.0)	906.0 (318.0 - 1377.0)	523.8 (278.0 - 658.0)

Table 1. Arthroscopic experience according to surgical logbooks of participants. Figures reported as mean (minimum – maximum), rounded to one decimal point.

	Participant experience level			Pairwise testing	
	Novice	Trainee	Expert	Novice vs. Trainee†	Trainee vs. Expert†
Diagnostic Time (seconds)	419.8±217.6 (145.7 - 920.7)	146.6±66.1 (72.0 - 381.2)	69.7±2.4 (66.8 - 72.5)	0.0001	0.0018
Diagnostic Hand Movements	338.9±183.1 (86.0 - 654.0)	137.0±63.7 (67.0 - 372.0)	68.0±12.7 (57.0 - 83.0)	0.0001	0.0047
Diagnostic Path Length (centimetres)	3139.8±1538.9 (1053.0 - 6122.9)	1336.7±539.4 (364.9 - 2837.3)	660.8±91.8 (578.1 - 782.8)	0.0001	0.0056
Diagnostic GRS	12.5 (10)	25 (5)	29 (0.8)	0.0001	0.0064
Bankart Time	472.9±48.1 (343.7 - 493.7)	318.4±69.0 (205.1 - 464.6)	217.5±65.0 (161.8 - 300.9)	0.0001	0.0251
Bankart Hand Movements	446.96±35.7 (330.0 - 461.0)	386.3±81.9 (268.0 - 640.0)	280.8±66.2 (218.0 - 374.0)	0.003	0.0292
Bankart Path length	5063.7±713.2 (3190.8 - 5452.4)	4062.8±932.6 (2088.0 - 7314.2)	3114.4±556.5 (2636.1 - 3906.3)	0.0003	0.0251
Bankart GRS	6 (7.5)	18 (5)	29 (4.3)	0.0001	0.0015

Table 2. Performance results for arthroscopic tasks grouped by experience level. Time taken, hand movements and path length reported as mean±standard deviation (minimum – maximum). GRS reported as median (interquartile range). Figures rounded to one decimal place. † p-value of Mann-Whitney U test.

Spearman rank correlation	**Diagnostic shoulder GRS**	**Bankart GRS**

coefficient (r_s)			
Number of arthroscopies performed	r_s value	0.7683	0.7698
	p value	< 0.0001	< 0.0001

Table 3. Correlation between number of previous arthroscopies performed and GRS on the two simulated tasks.

	Diagnostic task		**Bankart task**	
	Number of procedures	AUC	Number of procedures	AUC
Competent level of performance	52 95% CI (2 to 102) sensitivity 72.2% specificity 66.7%	0.82	248 95% CI (38 to 458) sensitivity 100% specificity 80%	0.96
Expert level of performance	290 95% CI (126 to 454) sensitivity 100% specificity 79.1%	0.86	476 95% CI (293 to 659) sensitivity 100% specificity 90.1%	0.94

Table 4. Optimum cut-off points of ROC curves to estimate number of procedures to reach a given level of performance, with area under the curve (AUC) shown for each analysis. (CI = Confidence interval)