Introduction

One of the challenges of competency-based graduate medical education is identifying accurate metrics for assessment of curricular activities. In surgical education, this includes evaluating the technical skill of trainees. In general surgery, assessment of technical skill using global rating scales has been found to be valid and reliable.2, However, similar scales intended to evaluate arthroscopic skill have not been widely accepted due to their limited ability to demonstrate valid and reliable assessment.3-5. The Arthroscopic Surgery Skill Evaluation Tool (ASSET) was developed as a global assessment arthroscopic technical skill with the potential to overcome the limitations of prior assessment scales (Figure 1). The ASSET was designed to be valid, reliable, simple to use, cost effective, and generalizable to multiple procedures and environments (simulation lab/operating room). The purpose of this project was to rigorously test the validity and reliability of using the ASSET to assess performance of diagnostic arthroscopy in the simulation lab setting. It was hypothesized that this new assessment tool would provide an objective, valid, and reliable method for assessing arthroscopic proficiency.

Methods

Content Validation

A modified Delphi process was used to determine the consensus opinion of a national panel of 8 orthopedic educators from 7 institutions. Group members reached consensus on the domains to include in the ASSET using multiple rounds of web-based surveys and conference calls. The group also created checklists of suggested tasks to be performed when using the ASSET to assess diagnostic arthroscopy of the knee and glenohumeral joints. It is intended that further task checklists can be created enabling ASSET to be used to assess multiple arthroscopic procedures.

Validity and Reliability in the Simulation Lab

Twenty-eight orthopedic residents and 2 orthopaedic faculty members with fellowship training in sports medicine performed diagnostic arthroscopy of a left and right cadaver knee in our simulation lab. Intraoperative video was recorded using the arthroscopic tower and included only the intra-articular view of the joint. All video recordings (n=60) were assigned a random identification number and reviewed by 2 raters blind to subject identity. Raters independently assigned scores to each video recording using the ASSET (Figure 1).

Results

Content Validation

The expert panel identified 8 global domains that could be used to assess multiple arthroscopic procedures in both the operating room and simulation lab (Figure 1).

Validity of the ASSET

Single factor analysis of variance demonstrated that there was a significant difference in mean ASSET scores assigned to subjects in the novice (PGY 1-2), intermediate (PGY 3-4), and advanced (PGY 5 and Attending) levels of training demonstrating construct validity (Figure 2). ROC curve analysis identified 50 cases as the number of prior knee arthroscopies most likely to predict attaining a passing score on the ASSET (Figure 3). This value was similar to a previously reported estimate of the number of cases needed for the average resident to attain proficiency and demonstrated the validity of using the ASSET as a pass/fail assessment.7

Reliability of the ASSET

Cronbach’s α demonstrated that the domains of the ASSET had good internal consistency for both raters suggesting that all domains were measuring the same construct (r=0.94). There was no significant difference in mean ASSET score assigned by each rater (p=0.93) and there was a strong positive correlation between the raters’ scores (Figure 4).

Conclusions

The ASSET was shown to be a valid and reliable method for assessing technical skill performing diagnostic arthroscopy in the simulation lab. The results demonstrate the ability of the ASSET to withstand rigorous testing and produce psychometric properties sufficient for high stakes examination. Further investigation is ongoing to determine the generalizability of using the ASSET to assess other procedures in the simulation lab and operating room. The development of an efficient way to train raters to use the ASSET is also underway.

References