INTRODUCTION

Progress in surgical education and mechanical skill can be difficult to measure due to the complexity of surgical procedures, the need for supervision and attending participation, and the need to provide for patient safety. Flexible laryngoscopic evaluation is commonly used for routine evaluation of the upper aerodigestive tract, and is commonly used by residents in the inpatient setting to evaluate the airway of patients presenting with acute complaints. Findings on these procedures range from normal to critical with immediate consequences for management of the patient’s airway or diagnosis of malignancy, and to date there are no objective measurements of milestones in this procedure.

MATERIALS AND METHODS

Twenty-one videos were created from archived endoscopy videos taken during routine clinical evaluations in the Otolaryngology outpatient clinic. Videos were specifically selected by the authors for presence of certain types of pathology or lack thereof. The videos were de-identified, and then the videos were edited to short representative examinations of 20-50 seconds. All videos were administered without sound so as not to bias the viewer toward the diagnosis because recognition of the examiner’s voice might bias the viewer toward a specific diagnosis.

Participants for the study were chosen to provide representative levels of training. The videos were administered at the start of the academic year, and the videos were shown to residents from each training level from incoming PGY-1 through outgoing PGY-5s (six classes of residents), and three attending Otolaryngologists. Of the attending Otolaryngologists, one was a fellowship trained rhinologist, and one was a general Otolaryngologist. A standardized questionnaire was given to each participant for each video that they viewed, totaling 25 questionnaires for each participant (see Figure 1). The information from the questionnaires was then recorded in a spreadsheet. Data analysis was performed using Statistical Analysis Software (SAS) and Excel. Intra-class correlation coefficients were calculated to determine the degree of agreement between observers of the same training level (ie PGY-1, PGY-2, etc) for each question asked on the questionnaire. Intra-class correlation coefficients were also calculated comparing the composite scores for each level of training against the composite attending “gold standard” scores.

RESULTS

Intra-class correlation coefficients were generated to compare the degree of agreement within each class of training (incoming PGY-1 through outgoing PGY-5, and attending Otolaryngologists) as well as the degree of agreement between each class of trainee against the attending Otolaryngologist gold standard. For interpretation, we concluded that 1.0-0.8 was very good agreement, 0.8-0.6 was good agreement, 0.6-0.4 was moderate agreement, 0.4-0.2 was fair agreement, and 0.2-0.0 was poor agreement. The intra-class correlations within each level of training demonstrated a large variations in reliability (See Table 1).

There were several categories with notably higher intra-class correlations when comparing participants within their own class of training, including subglottic stenosis, left true vocal cord immobility, right true vocal cord immobility, and level of concern. When comparing intra-class correlations between different levels of training and the attending Otolaryngologist gold standard, the authors found that there were large variations in reliability but that the same categories demonstrated higher reliability, subglottic stenosis, left true vocal cord immobility, right true vocal cord immobility, and level of concern (See Table 2).

CONCLUSIONS

The authors provide evidence that resident competency with interpretation of flexible nasolaryngoscopy can be measured objectively by comparing resident interpretation to that of attending Otolaryngologists. The best measurements of competency are in aspects of the examination that demonstrate increasing interpreter agreement with level or training. These specific aspects are vocal cord mobility, subglottic stenosis, and level of concern of airway patency. Resident operators approach attending levels of reliability between the PGY2 and PGY 3 years.