360 - Degree Evaluations of Senior EM Residents: A Necessary Evil

Gregory Garra, DO • Andrew Wackett, MD • Henry Thode, PhD
Department of Emergency Medicine, Stony Brook University, Stony Brook, NY

Background:
- The ACGME recommends multisource feedback (MSF) evaluation of resident performance.
- The utility of MSF for EM trainees is untested.

Objective:
- To determine the feasibility of an MSF program and evaluate the intra and interclass correlation of a previously reported resident professionalism evaluation - The Humanism Scale (HS).

Methods:

Study Design
- Survey

Subjects
- Ten EM - 3 residents

Measures
- An anonymous 9-item modified HS (EM-HS) was distributed to the ED nursing staff, attending physicians and patients (figure 1). Patient surveys consisted of questions 4-9.
- The evaluators rated resident performance on a 1-9 scale (needs improvement to outstanding).
- Residents were asked to complete a self evaluation of performance using the same scale.

Analysis
- Generalizability coefficients were used to assess the reliability within evaluator classes.
- The mean score for each of the 9 questions provided by each evaluator class was calculated for each EM resident.
- Correlation coefficients were used to evaluate correlation between rater classes for each question on the EM-HS.
- Generalizability coefficients and correlations > 0.70 were deemed acceptable.

Results:
- EM - HS’s were obtained from 44 nurses and 12 attending physicians.
- Each resident had an average of 18 evaluations by ED patients at the point of care.
- Reliability within evaluator class was acceptable: Ep 0.79, 0.83 and 0.77 for attendings, nurses and patients, respectively.
- Inter-class reliability was poor for all pairs of evaluator types (table 1).

Limitations:
- Single-center study
- Modified survey
- Memory biases such as context effect, mood congruent memory bias and distinctive encoding
- Not designed to assess validation

Conclusions:
- Intraclass reliability was acceptable for all evaluator types.
- Ratings were not consistent across rater classes, confirming the utility of MSF instruments.