Reliability and Validity Testing of a Rubric to Assess Medical Student Proficiency for Informed Consent

Suzanne G. Wilson, Kirsten Waara, Faith Palmer, Elizabeth Petsche, Leana Danesh, Brandy Church, Samuel Wisniewski, William Corser and John McDaniel

Objectives

- Evaluate whether the draft rubric is a reliable and valid tool for assessing medical student proficiency for providing informed consent ("IC").
- Assess whether ratings submitted by the three planned evaluator groups (medical students, residents and clinical faculty) are comparable.
- Incorporate revised rubric into the preclerkship Ethics Course.

Background

Informed consent is one of the entrustable professional activities ("EPAs") that medical students are expected to be able to independently perform prior to graduation. The goal of this project was to conduct reliability and validity testing of a proposed rubric to assess medical student performance of verbal informed consent. Knowledge of informed consent is taught and assessed in the year 2 Ethics course at MSUCOM, however skill evaluation had yet to be defined. This project aimed to fill that gap, creating and testing a rubric to be used for assessment of medical student skill for providing informed consent. It also hoped to appraise how peer and resident evaluations compared to faculty for if comparable to faculty ratings, these might provide additional useful input in the entrustment decision for this EPA.

Educational Design/Methods

Review of the literature determined criteria for inclusion on the draft rubric. Following MSU IRB approval as an exempt study, an emailed consent form was sent to eligible subjects. Those who wished to participate indicated their consent to participate by reply email. Participants were sent a copy of the rubric, a link to view the four-minute sample videotaped informed consent encounter online and a link to complete the rubric online via MSU Qualtrics. Participants all watched the same informed consent video and completed the online rubric via MSU Qualtrics. For statistical analysis, determination had to be made about whether each criteria of introducing self to patient and whether the information was presented in a logical sequence. Four criteria (shown in red on Figure 1 below) had high variability, indicating a need for revision to improve clarity or possible elimination from the tool. With an odds ratio of 7.5, clinical faculty were more likely than residents to correctly appraise whether the purpose of IC had been addressed (95% Confidence Intervals). With an odds ratio of 12.8, clinical faculty were more likely than students to correctly appraise if the uncertainties of the procedure had been addressed (95% Confidence Intervals).

Data Collection/Analysis

MSU Qualtrics data was exported for statistical analysis with support from Dr. Corser, Samuel Wisniewski and Student Doctor Palmer. Recruitment fell short of goals, yielding only 40 participants: 13 medical students, 7 residents and 20 clinical faculty. The shortfall of participants limited the options for statistical analysis. It was also noted that the use of a single videotaped encounter resulted in an effect size of 1, which also limited techniques that could be used for statistical analysis.

Outcomes / Results

- There were four criteria that all participants answered correctly (patient circumstances, patient questions/concerns, indications for procedure, and speaking in a clear/audible/unhurried manner).
- On average, all 3 groups performed similarly overall (72-78% correct).
- Significant difference (p = .004) noted between groups on criterion of discussing uncertainties associated with the procedure.
- All groups performed poorly (<30% correct) on the criteria of introducing self to patient and whether the information was presented in a logical sequence.
- Four criteria (shown in red on Figure 1 below) had high variability, indicating a need for revision to improve clarity or possible elimination from the tool.
- With an odds ratio of 7.5, clinical faculty were more likely than residents to correctly appraise whether the purpose of IC had been addressed (95% Confidence Intervals).
- With an odds ratio of 12.8, clinical faculty were more likely than students to correctly appraise if the uncertainties of the procedure had been addressed (95% Confidence Intervals).

Discussions

- Overall, ratings of clinical faculty tended to be more correct than those of students, an expected finding based on their greater experience levels.
- Ratings of students and residents were comparable.
- Clinical faculty were more precise with their analysis and answers, such as being critical of individualizing conversation, perhaps due to having more basis for recognizing when this was not being done.
- Students had difficulty identifying adequacy of discussion of risks and complications, possibly due to their limited background.
- All groups varied in opinion whether the sample encounter was presented in a logical sequence. Logical flow is a subjective judgment of individual clinicians at present.
- It was surprising that even criteria that seemed clear, such as "introduced self" displayed variability as assessed by the cohort groups.
- It could be inferred that clinical faculty understood the criteria more clearly because they regularly deliver IC in provision of care.

Conclusions and Implications

- Use of a literature basis for rubric construction was helpful in establishing content validity and for identifying specific criteria to be included.
- Participant comments supported the utility of the rubric and validity of its component criteria for evaluating this EPA.
- Four high variability criteria were modified for greater clarity and the revised form was used in a learning activity in the Ethics course. A replication study using the revised rubric and a larger sample size is indicated.
- Logical flow of IC is very subjective. Future research could explore what sequence of topics is most beneficial to promote patient understanding so that this can be standardized in the training of medical students.
- Because this study found faculty ratings differed from those of students/residents, this may support using faculty for IC assessment if it is for high stakes (a grade) and not just formative feedback.
- Following the study, a revised rubric was used in a peer feedback exercise in the Ethics course and was deemed a useful tool for formative assessment.

References

Core EPAs guiding principles. Association of American Medical Colleges.