Measurement of medical student communication skills during real patient encounters compared to communication skills measured by OSCE’s

MARY LACY, MD
LEONARD NORONHA, MD
YURI LEYVA, MS
RUSH PIERCE, MD

SCHOOL OF MEDICINE
Objectives

Describe the importance of communication skills for patient care

Describe some of the issues with summative assessment of communication skills using standardized patients

Describe, compare, and contrast the Direct Observation of a Clinical Encounter (DOCE) and the Objective Structured Clinical Examination (OSCE) experiences at UNM
Effective provider-patient communication can improve health outcomes

Review – 21 studies (RCTs, cohort studies)
- 16 showed positive results

Communication in history-taking or discussion of management plans

Emotional health
Symptom resolution
Function
Physiologic measures

Effective communication skills can be taught and are learned

Experiential training
Should be taught *within* clinical clerkships

Feedback
Should be given to all students

DOCE

- Formative with feedback
- Single Encounter in Longitudinal Relationship
- Inpatient setting
- No time limit (though it is timed)
- Scoring by MD preceptor (1 observation)
OSCEs may not reliably assess communication skills

Meta-analysis: 39 studies, 188 alpha values

Across station scores:

Overall alpha = 0.66 (0.62 – 0.70)
Clinical skills alpha = 0.69 (0.66 – 0.73)
Communication scales = 0.55 (0.45 – 0.63)

“It appears to be more difficult to reliably assess communication skills than clinical skills across stations.”

Medical Education 2011: 45: 1181–1189
Study Question: Are communication skills measured by faculty using a modified NM-CCS similar to skills measured in the OSCE?

<table>
<thead>
<tr>
<th>DOCE MODIFIED NM-CCS</th>
<th>OSCE NM-CCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Open the discussion</td>
<td>1. Open the discussion</td>
</tr>
<tr>
<td>2. Build a relationship</td>
<td>2. Build a relationship</td>
</tr>
<tr>
<td>3. Gather information</td>
<td>3. Gather information</td>
</tr>
<tr>
<td>4. Understand the Patient’s Perspective</td>
<td>4. Understand the Patient’s Perspective</td>
</tr>
<tr>
<td>5. Share Information</td>
<td>5. Share Information</td>
</tr>
<tr>
<td>6. Provide Closure</td>
<td>6. Reach an Agreement</td>
</tr>
<tr>
<td>7. Provide Closure</td>
<td>7. Provide Closure</td>
</tr>
</tbody>
</table>

Total score: 6 - 30

Total Score: 7 - 35
Comparison specific to UNM DOCE and OSCE

**DOCE**
- Formative with feedback
- Single Encounter in Longitudinal Relationship
- Inpatient setting
- No time limit (though it is timed)
- Scoring by MD preceptor (1 observation)

**OSCE**
- Summative
- Standardized patients
- Single encounter
- Typically outpatient setting
- 15 minute time limit
- Scoring by SP (5 cases)

**Student observed on communication and examination skills**

**NM-CCS**
measured by faculty using a modified NM-CCS similar to skills measured in the OSCE?

<table>
<thead>
<tr>
<th>6. REACH AGREEMENT (Planning Evaluation and Treatment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Negotiation</strong></td>
</tr>
<tr>
<td>○ No plan or a coercive plan</td>
</tr>
<tr>
<td>○ Ignores your ideas/requests without explanation</td>
</tr>
<tr>
<td>○ Requests feedback on the plan presented</td>
</tr>
<tr>
<td>○ Verifies your understanding of the final plan and requests your feedback</td>
</tr>
</tbody>
</table>
Mean OSCE and DOCE scores were similar. Students tended to perform better in Opening on OSCE and Closing on DOCE.

<table>
<thead>
<tr>
<th></th>
<th>DOCE %, mean (90% C.I.)</th>
<th>OSCE %, mean (90% C.I.)</th>
<th>Matched p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total score</strong></td>
<td>60 (59, 61)</td>
<td>60 (59, 61)</td>
<td>0.9431</td>
</tr>
<tr>
<td><strong>1. Open the Discussion</strong></td>
<td>54 (52, 56)</td>
<td>60 (59, 61)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>2. Build a relationship</strong></td>
<td>63 (62, 65)</td>
<td>61 (59, 62)</td>
<td>0.1051</td>
</tr>
<tr>
<td><strong>3. Gather information</strong></td>
<td>60 (58, 62)</td>
<td>60 (59, 61)</td>
<td>1.0000</td>
</tr>
<tr>
<td><strong>4. Patient’s Perspective</strong></td>
<td>60 (58, 62)</td>
<td>61 (59, 62)</td>
<td>0.5747</td>
</tr>
<tr>
<td><strong>5. Share Information</strong></td>
<td>59 (58, 61)</td>
<td>61 (60, 63)</td>
<td>0.1028</td>
</tr>
<tr>
<td><strong>6. Reach an Agreement</strong></td>
<td>NA</td>
<td>60 (58, 63)</td>
<td></td>
</tr>
<tr>
<td><strong>7. Provide Closure</strong></td>
<td>65 (62, 67)</td>
<td>58 (55, 60)</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

4 observers
83 observations
* Two sample T-test
The mean score on DOCE was not statistically significantly different between groups who passed or failed the OSCE

<table>
<thead>
<tr>
<th></th>
<th>Students who fail OSCE (n=26)</th>
<th>Students who pass OSCE (n=57)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean DOCE % score† (SD)</td>
<td>60.3 (6.4)</td>
<td>59.9 (6.1)</td>
<td>0.8296</td>
</tr>
</tbody>
</table>

OSCE Pass Score >/= 21 (60%)  
† Two-sample independent t-test
While mean total scores were similar, paired profiles are not correlated.
While mean total scores were similar, paired profiles are not correlated.
Total scores from individual preceptors were not statistically significantly different.
Time spent at bedside on DOCE was not a significant predictor for DOCE or OSCE scores.

**Fit Plot for DOCE**

- $R^2 = 0.0166$
- $p = 0.2464$

**Fit Plot for OSCE**

- $R^2 = 0.01$
- $p = 0.3687$
Faculty observing 3rd year students in a standardized bedside observation in the hospital score communication skills differently than when the same students are observed by SP’s in an OSCE.

Some Questions:

Do the limitations of this study preclude our ability to make any judgement on our study question?

Would students who are discordant in performance on OSCE vs DOCE benefit from additional communication skills training?

Is physician communication style and skillfulness different in different clinical settings?
  ◦ If so, should our OSCE’s be designed with various clinical settings?

What level of reliability is acceptable for a high-stakes OSCE or Step-2 CS?