The Ottawa Surgical Competence Operating Room Evaluation
O-SCORE
A Tool to Assess Surgical Competency

Wade T. Gofton MD MEd FRCSC
Nancy L. Dudek MD MEd FRCPC
Timothy J. Wood PhD
Fady Balaa MD MEd FRCSC
Stanley J. Hamstra PhD
Disclaimers

- Gofton
  - DePuy-Synthes
    - Institutional Research Support
  - Zimmer
    - Implant Design Consultant
  - Smith and Nephew
    - COTS research study
Background

- Certification bodies assess competence based on knowledge, not *actual ability to perform*

- Responsibility of residency programs
  - Technical performance poorly assessed
• Experts know what competent clinical performance entails

• Surgical mini-CEX
  – Multiple assessments
  – Different raters
Objectives

• Develop a succinct surgical assessment tool that could be used for any surgical procedure

• Change the focus of assessment away from performance relative to year of training to competent performance
Methods

• Phase 1: Developing, Piloting and Refining the O-SCORE

  – Expert group developed a 12 question scale
  – Piloted in the Division of Orthopedic Surgery
    • Oct 2008- Feb 2009
  – Psychometric analysis
    • Internal structure (descriptive stats, correlations across items, generalizability analysis, comparisons across PGY years)
  – Qualitative assessment
    • Response process and feasibility
  – Revised to a 9 point scale
The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE)

<table>
<thead>
<tr>
<th>Trainee #:</th>
<th>Level: 1 2 3 4 5</th>
<th>Staff:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure:</td>
<td></td>
<td>Date:</td>
</tr>
</tbody>
</table>

Relative complexity of this procedure to average of same procedure

- **Low**
- **Medium**
- **High**

The purpose of this scale is to evaluate the trainee’s ability to perform this procedure safely and independently. With that in mind please use the scale below to evaluate each item, irrespective of the resident's level of training in regards to this case.

**Scale**

1. **Preprocedure plan**
   - Gathers/assesses required information to reach diagnosis and determine correct procedure required

2. **Case preparation**
   - Patient correctly prepared and positioned, understands approach and required instruments, prepared to deal with probable complications

3. **Knowledge of specific procedural steps**
   - Understands steps of procedure, potential risks, and means to avoid/overcome them

4. **Technical performance**
   - Efficiently performs steps, avoiding pitfalls and respecting soft tissues

5. **Visuospatial skills**
   - 3D spatial orientation and able to position instruments/hardware where intended

6. **Postprocedure plan**
   - Appropriately complete post procedure plan

7. **Efficiency and flow**
   - Obvious planned course of procedure with economy of movement and flow

8. **Communication**
   - Professional and effective communication/utilization of staff

9. **Resident is able to safely perform this procedure independently (circle)**
   - Y
   - N

10. **Give at least 1 specific aspect of procedure done well**

11. **Give at least 1 specific suggestion for improvement**

Signatures: Staff:

Trainee:
Scale
1 – “I had to do” – i.e. Requires complete hands on guidance, did not do, or not given the opportunity to do
2 – “I had to talk them through” – i.e. Able to perform tasks but requires constant direction
3 – “I had to prompt them from time to time” – i.e. Demonstrates some independence, but requires intermittent direction
4 – “I needed to be in the room just in case” – i.e. Independence but unaware of risks and still requires supervision for safe practice
5 – “I did not need to be there” – i.e. Complete independence, understands risks and performs safely, practice ready

1. Pre-procedure plan
   Gather/assesses required information to reach diagnosis and determine correct procedure required

2. Case Preparation
   Patient correctly prepared and positioned, understands approach and required instruments, prepared to deal with probable complications

3. Knowledge of specific procedural steps
   Understands steps of procedure, potential risks and means to avoid/overcome them

4. Technical Performance
   Efficiently performs steps avoiding pitfalls and respecting soft tissues

5. Visuospatial Skills
   3D spatial orientation and able to position instruments/hardware where intended

6. Post-procedure plan
   Appropriate complete post-procedure plan

7. Efficiency and Flow
   Obvious planned course of procedure with economy of movement and flow

8. Communication
   Professional and effective communication/utilization of staff
9. Resident is able to safely perform **this** procedure **independently** (circle)  
   
   Y   N

10. Give at least 1 **specific** aspect of procedure done well

11. Give at least 1 **specific** suggestion for improvement

Signatures:  Staff:

Trainee:
Methods

• Phase 2: Testing the O-SCORE across specialties
  - OHREB
  - Orthopedic and General Surgery
    – January 2011 – May 2011

  – Orthopedic Surgery
    • 6 procedures

  – General Surgery
    • 5 procedures

  – Similar quantitative and qualitative analysis
Results

- Phase 1: Piloting and Refining the OSCORE
  - 20 ortho residents with 72 procedures evaluated
  - 9 observations excluded (incomplete)
  - Ave 3.15 observations/resident
  - Generalizability = high reliability (0.82)
  - High correlation between items (0.68-0.86)
  - Significant effect of PGY level, all levels differing except PGY 4-5
  - Trainees felt it improved feedback
### Phase 2: Testing the O-SCORE across specialties (internal structure)

- 34 surgeons, 37 residents, 163 procedures
- Ave 4.41 observations per resident (range 1-25)

<table>
<thead>
<tr>
<th>O-SCORE item</th>
<th>Rating(^{\dagger})</th>
<th>Range(^{\dagger})</th>
<th>Corrected item–total correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
</tr>
<tr>
<td>1. Preprocedure plan</td>
<td>4.24</td>
<td>1.01</td>
<td>1</td>
</tr>
<tr>
<td>2. Case preparation</td>
<td>4.21</td>
<td>0.85</td>
<td>1</td>
</tr>
<tr>
<td>3. Knowledge of specific procedural steps</td>
<td>3.90</td>
<td>0.93</td>
<td>1</td>
</tr>
<tr>
<td>4. Technical performance</td>
<td>3.61</td>
<td>1.01</td>
<td>1</td>
</tr>
<tr>
<td>5. Visuospatial skills</td>
<td>3.94</td>
<td>0.97</td>
<td>2</td>
</tr>
<tr>
<td>6. Postprocedure plan</td>
<td>4.47</td>
<td>0.75</td>
<td>2</td>
</tr>
<tr>
<td>7. Efficiency and flow</td>
<td>3.72</td>
<td>0.99</td>
<td>1</td>
</tr>
<tr>
<td>8. Communication</td>
<td>4.31</td>
<td>0.75</td>
<td>2</td>
</tr>
</tbody>
</table>
• Reliability
  – Differences between trainees and within each trainee account for the largest variability
  – Specialty did not account for much variability

• G coefficient 0.80

• Would take at least 5 observations per trainee
• Validity
  – Mean Total procedure scores compared with safe to perform independently (question 9)

  • Not ready to safely perform
    – 3.58 ± 0.61, n=84 observations

  • Ready to safely perform
    – 4.54 ± 0.48, n=79 observations

  – Correlation 0.66 (p < .001)
• Validity
  – Mean total procedure scores increased with PGY year
    – PGY 1&2 (n=44) – 3.57±0.7
    – PGY 3 (n=61) – 4.08±0.61
    – PGY 4&5 (n=33) – 4.38±0.68
    \[ P < .001 \]
    \[ P < .02 \]
– Qualitative assessment (consequences)
  • Practical and useful for staff and residents

  • Anchors made it easier to evaluate
    – did not have to decide if above or below PGY level

  • The ‘colloquial’ anchors reflected how they determined a trainee’s level of independence
    – supporting response process and providing evidence for validity of the O-SCORE

  • Residents accepted low scores, pleased with improved feedback
• Kane and Messick
  – Validity requires an evidentiary chain demonstrating that it is measuring the desired construct, evidence is collected from several sources

  – Content
    • Relationship between tests content and the construct it is intending to measure
      – lit search, expert group
  – Response Process
    • Fit between the construct and the detailed nature of performance engaged in
      – Focus group
      – Tool aided in directing staff and resident to desired behavior and well accepted
– Internal structure
  • Reliability and Factor Analysis
    – Quantitative analysis consistent in both phases of study and between programs

– Relation to other variables
  • Correlation with other instruments or outcomes

– Consequences
  • Evaluation of intended or unintended consequences
    – qualitative analysis (focus groups)
Limitations

• Form of direct observation without criteria
  – ? Unreliable
  – Relies on experts

• Unblinded raters
Uses

• Improve reliability of ITER

• Competency-based procedure log
  – Demonstrate competency
  – Early warning system
  – Aid in program design

• Just one more tool to aid in a competency based assessment
Conclusions

• O-SCORE able to accurately differentiate between junior, mid-level and senior residents

• With most procedures a trend to increasing competence with PGY level

• Effective in both surgical specialties across multiple procedures

• Accepted by staff and residents
Support

• Pilot Study
  – Royal College of Physicians and Surgeons of Canada Medical Education grant

• Across Specialties
  – AIME grant
  – PSI grant
Thanks

- General and Orthopedic Surgery Residents and staff
- Shey Seth
- Julia Foxall