The Optimization and Standardization of the UEMS Section of Surgery Exams

Division of HPB in association with E-AHPBA
Pilot Exam Proposal in HPB

Mr Donal O’Connor
Prof Kevin Conlon
Prof Xavier Rogiers
European Union of Medical Specialties

• Why are EBSQ important?

• Heterogeneity exists across member states in:
  – Post graduate specialist training
  – Assessment and final examination standards

• It is vital that high stakes competency assessments are reputable
European Union of Medical Specialties

Examinations in Section of Surgery

• Surgical assessments have been criticised for lacking
  – Transparency
  – Reliability
  – Validity

• Within HPB division there were concerns about the exam
  – Exam not formally mapped to curriculum
  – No 360° evaluation
  – Exam validity not assessed
  – Examiner reliability not assessed

• Do similar limitations exist in other division exams?

In reviewing the HPB exam we saw an Opportunity:
Survey all Heads of Divisions that deliver exams

Collate the strengths (and weakness) of their respective exams

Perform qualitative assessment of exams
  – Survey of examiners and candidates from recent exam diets

Determine the level of ‘commonality’ for standardized assessment and evaluation methods
European Union of Medical Specialties

Proposal: UEMS-HPB Exam Project

- Publish survey and blueprint for the ‘ideal’ exam format
- Deliver an exam using HPB as a prototype
- Validate: Qualitative and Quantitative, Internal and External
- Make the blueprint transferable to other specialties
- Reproduce and publish methods
The Ideal Assessment

‘No single assessment method can provide all the data required for the judgment of anything so complex as the delivery of professional services by a successful physician’

George Miller
Proposal: UEMS-HPB Exam Project

The Written exam

- MCQ
  - Single best answer
  - True / false

- Questions provided by E-AHPBA members

- Questions reviewed by an examination board and mapped to curriculum

- Question bank developed to assess basic knowledge & clinical reasoning skills
  - Script Concordance test\(^1\)
  - Internal (E-AHPBA) and External (IHPBA) expert review for validation and correlation

- Minimum standard to be determined by established methods (Angoff\(^2\))

---

2. Vehoeven BH et al. Reliability and credibility of an Angoff standard setting procedure in progress testing using recent graduates as judges. Medical Education. 1999
Sample Questions

Question 2

A 71 year old man, 9 days following a pancreaticoduodenectomy, presents a drain output of 100cc/day of water like fluid with 15000 IU of amylase content. The patient is in good clinical condition and the blood examination shows low inflammatory parameters. Antibiotics and pancreatic exocrine antisecretory drugs are used as well as total parenteral nutrition. My diagnosis is:

A. Delayed gastric emptying
B. Pancreatic Fistula Grade C
C. Pancreatic Fistula Grade B
D. Pancreatic Abscess

Question 29

A 36-year-old woman was treated with a distal pancreatectomy (Warshaw procedure) for a pancreatic tail lesion. Two days after the procedure she complained about severe pain mainly situated in the left upper part of her abdomen. A CT scan was made which shows:

A. Delayed gastric emptying
B. Splenic infarction
C. Subphrenic fluid collection
D. Liver vein thrombosis

- Collaborate with Umbrella Consortium for Assessment Networks (UCAN)
- Item management system (IMS): Secure, web-based tool for creation and statistical analysis of exam questions
- Suitable for web-based exam delivery
European Union of Medical Specialties

Proposal: UEMS-HPB Exam Project

The Oral Exam

- Allow assessment of high levels of clinical competence and professionalism\textsuperscript{1,2}
- Demonstrate classic Bayesian decision theory\textsuperscript{3}

2 Kayden JB. Bayesian methods for health related decision making. Statistics in Medicine. 2005
Guided oral exam

- Examiners provided with structured cases & suggested questions
- Marking guidelines provided\(^1,2\)
- Complex scenarios can be discussed to assess integrated professional judgment

1 Luxenberger W et al. The development and design of the European Board of Otorhinolaryngology-Head and Neck Surgery Examination. Eur Arch Otorhinolaryngol. 2016
European Union of Medical Specialties
Proposal: UEMS-HPB Exam Project

• Pilot: Dublin December 2015
• Guided oral exam

• 3 stations - 30 minutes each
  – Pancreas
  – Liver
  – Gallbladder and Bile duct

• Examiners provided with structured cases, suggested questions, marking scheme
• iPADs used to display clinical/radiology images
• Candidates examined by 2 different examiners at each station
• Examiners meeting before and after oral exam
Example of Clinical Cases on Tablet

History

- 35-year-old male patient
- Chronic pancreatitis diagnosed 5 years
- Diabetic and severe exocrine deficiency
- Abdominal pain requiring opiates
- CA 19-9 normal
- Medical history:
  - Pancreatic duct stenting for chronic pancreatitis

Suggested Questions

1. Describe the radiological pictures
2. What is your diagnosis?
3. What is the treatment?

Answers & Discussion

1. Pictured to be described (CT and MRI)
2. Diagnosis: Chronic pancreatitis
   a. Enlarged head of pancreas on CT with dilated duct and calcifications
   b. Dilated tortuous pancreatic duct on MRI
3. Treatment: Consider surgery
   a. Discuss nutritional management in chronic pancreatitis
   b. Discuss pain management and endoscopic PD stenting
   c. Discuss surgical drainage and resectional procedures
Validating the Exam: Qualitative

• Internal and external review of cases for clinical relevance

• Examiners meeting and feedback
  Control the ‘hawk and dove’ effect

• Survey of examiners and candidates

1 McManus et al. BMC Medical Education.2013
### TABLE 5. Interrater Reliability of the “ACE” Format

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach α</td>
<td>0.907</td>
<td>0.766–1</td>
</tr>
<tr>
<td>Krippendorff α</td>
<td>0.824</td>
<td>0.824–1</td>
</tr>
<tr>
<td>Cohen κ</td>
<td>0.823</td>
<td>0.571–1</td>
</tr>
</tbody>
</table>

### Consistency of examiners

**Clinical & Written Analysis: Consistency of performance**

<table>
<thead>
<tr>
<th>Correlations</th>
<th>ENT</th>
<th>OPTHO</th>
<th>P1TOTAL</th>
<th>P2TOTAL</th>
<th>LONG CASE</th>
<th>SHORT CASE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.034</td>
<td>.039</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>OPTHO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.288</td>
<td>.351</td>
<td>.502</td>
<td>.296</td>
<td>.314</td>
<td>.449</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>P1TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.260</td>
<td>.351</td>
<td>.777</td>
<td>.719</td>
<td>.708</td>
<td>.888</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>P2TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.307</td>
<td>.502</td>
<td>.777</td>
<td>.724</td>
<td>.685</td>
<td>.899</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>LONG CASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.177</td>
<td>.296</td>
<td>.719</td>
<td>.724</td>
<td>1</td>
<td>.692</td>
<td>.883</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.034</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>SHORT CASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.172</td>
<td>.314</td>
<td>.708</td>
<td>.685</td>
<td>.692</td>
<td>1</td>
<td>.869</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.039</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>.302</td>
<td>.449</td>
<td>.888</td>
<td>.899</td>
<td>.883</td>
<td>.869</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
<td>145</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).**

**. Correlation is significant at the 0.05 level (2-tailed).**

European Union of Medical Specialties

Proposal: UEMS-HPB Exam Project

• Next Phase
  – Examination in Maastricht June 2016
  – Complete the qualitative and quantitative analysis
  – Publish report of new exam format

• Evaluate survey of specialty divisions

• Create new model for UEMS examinations

• Determine transferable commonality of content e.g. written and oral format

• Reproduce and re-validate
European Union of Medical Specialties

Proposal: UEMS-HPB Exam Project

- Curriculum review (internal and external)

- Publish resources on website
  - Textbooks
  - Consensus guidelines and key journals
  - CME activities
  - Sample MCQ and clinical cases for formative assessment

- Preparatory postgraduate courses delivered by E-AHPBA
  - Map content to exam and desired professional specialty standards
  - Maastricht & Lyon 2016

- Develop a ‘Train the examiners’ course to hold concurrently with postgraduate courses
European Union of Medical Specialties

Surgeon in Training

Formative Knowledge Assessment
- Validated curriculum
- Sample MCQs and Cases
- Accredited courses
- Textbooks
- Consensus guidelines
- CME activities

Practical Skills Assessment
- e-Logbook
e-Portfolio

Summative Assessment of Clinical Competence and Specialization
- Validated Written and Oral Examinations

Specialist Surgeon

Adapted from Mills P et al. EBSC. 2008
European Union of Medical Specialties

Surgeon in Training

Formative Knowledge Assessment
- Validated curriculum
- Sample MCQs and Cases
- Accredited courses
- Textbooks
- Consensus guidelines
- CME activities

Practical Skills Assessment

Summative Assessment of Clinical Competence and Specialization

e-Logbook
e-Portfolio

Validated Written and Oral Examinations

Specialist Surgeon

Adapted from Mills P et al. EBSC. 2008
European Union of Medical Specialties

Surgeon in Training

Formative Knowledge Assessment
- Validated curriculum
- Sample MCQs and Cases
- Accredited courses
- Textbooks
- Consensus guidelines
- CME activities

Practical Skills Assessment
- e-Logbook
- e-Portfolio

Summative Assessment of Clinical Competence and Specialization
- Validated Written and Oral Examinations

Specialist Surgeon

Adapted from Mills P et al. EBSC. 2008
European Union of Medical Specialties

Conclusion

Benefit of a standardised UEMS exam

- Provide a validated & citable exam model acceptable to all specialty divisions

- Template of common features allows resources, research and feedback to be shared among divisions

- All divisions will benefit from aggregation of validation process

- Independence of specialty divisions and societies preserved: Exams will still feature diverse elements within specialty divisions

- Defendable against appeals and competition