Selecting International Graduates: Interpreting Medical Council of Canada Examination Results

SANDRA BANNER, IAN BOWMER, ANDRÉ-PHILIPPE BOULAIS, NATALIE AUGER

CARMS & MEDICAL COUNCIL OF CANADA

THE INTERNATIONAL CONFERENCE ON RESIDENCY EDUCATION

SEPTEMBER 2011
MCC Evaluating Examination Format

- Computerized since 2008
- 180 multiple-choice questions (30 pilot)
- 4 hours time to complete
- Offered up to 5 times a year in 2-3 week windows.
- Convenient scheduling, 500 sites in 80 countries.
- About 3000 candidates test each year.

- Prerequisite to CaRMs and the QE Part I
- Screening examination of basic medical knowledge:
  - Pediatrics
  - Psychiatry
  - Internal Medicine
  - Surgery
  - Obstetrics & Gynecology
  - Preventive medicine and community health.
All questions reflect the published MCC Objectives
  - http://www.mcc.ca/objectives_online

These include Considerations for Cultural-Communication, Legal & Ethical Aspects of Medicine, revised in 2009 to include
  - Professionalism
  - CanMEDS roles
### Evaluating Examination Content

#### Table

<table>
<thead>
<tr>
<th>Health Grouping</th>
<th>Total *Proportion of Questions</th>
<th>Proportion Data Gathering and Investigation</th>
<th>Proportion Patient Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>17</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Maternal</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Adult</td>
<td>50</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>Mental</td>
<td>16</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Population Health</td>
<td>9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>57</strong></td>
<td><strong>34</strong></td>
</tr>
</tbody>
</table>

* Average range is +/- two questions
Total scores are based on Item Response Theory 1-parameter model (i.e., difficulty).

Score reporting for the EE
- Pass mark is 250
- Average score is 271
- Standard deviation is 50
The passing score is set at 390 on the reporting score scale.
- Range of scores: 0-950, SD 100. Mean score is 500 for the spring cohort and 400 for the fall cohort.
- The passing score is established by expert judges using criterion-referenced procedures.

The proficiency definition is based on the required competence of a physician under supervised practice.
- A passing score was established in 2009 and is mapped forward for new forms. Revisited every 5 years.
In 2011, the program was successfully administered in Alberta, Manitoba, British Columbia, Ontario, and Québec.
### National Assessment Collaboration
**OSCE Content Guidelines**

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Systems</th>
<th>Quantity</th>
<th>Discipline</th>
<th>Quantity</th>
<th>Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least 1 (not more than 3 per system)</td>
<td>Respiratory</td>
<td>2-4</td>
<td>Medicine</td>
<td>2</td>
<td>Newborn, infant, pre-schooler or child.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(no more than four cases in medicine and surgery)</td>
<td></td>
<td>(male to female 40:60)</td>
<td></td>
</tr>
<tr>
<td>At least 1 (not more than 3 per system)</td>
<td>CV</td>
<td>2-4</td>
<td>Surgery</td>
<td>1</td>
<td>Adolescent 13-18</td>
</tr>
<tr>
<td>At least 1 (not more than 3 per system)</td>
<td>GI</td>
<td>1-3</td>
<td>Psychiatry</td>
<td>7-8</td>
<td>19-44, 45-64</td>
</tr>
<tr>
<td>2-3</td>
<td>MSK GU Endocrine Neurology</td>
<td>1-3</td>
<td>ObGyn</td>
<td>3-4</td>
<td>Over 65</td>
</tr>
<tr>
<td>2-3</td>
<td>Other</td>
<td>1-3</td>
<td>Pediatrics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Medical Council of Canada 2011
<table>
<thead>
<tr>
<th>Competencies Rated</th>
<th>CanMEDs Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>History Taking</td>
<td>Medical Expert</td>
</tr>
<tr>
<td>Physical Examination</td>
<td>Communicator</td>
</tr>
<tr>
<td>Differential Diagnosis</td>
<td>Collaborator</td>
</tr>
<tr>
<td>Investigations</td>
<td>Professional</td>
</tr>
<tr>
<td>Communication Skills</td>
<td>Health Advocate</td>
</tr>
<tr>
<td>Management</td>
<td></td>
</tr>
<tr>
<td>Language Fluency</td>
<td></td>
</tr>
<tr>
<td>Data Interpretation</td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td></td>
</tr>
<tr>
<td>Minimum 5 competencies are rated per station.</td>
<td>Used for tracking, not reported currently.</td>
</tr>
</tbody>
</table>

A holistic/global rating scale is used for each NAC OSCE station.
### TPx Content Guidelines

#### THERAPEUTICS

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Patient Group</th>
<th>Patient Group</th>
<th>Patient Group</th>
<th>Patient Group</th>
<th>Patient Group</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacotherapy (1)</td>
<td>Infants/Children (A)</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>14 (58%)</td>
</tr>
<tr>
<td>Adverse Effects (2)</td>
<td>Adolescents (13-17) (B)</td>
<td>*</td>
<td>2</td>
<td>1</td>
<td>*</td>
<td>4 (17%)</td>
</tr>
<tr>
<td>Disease Prevention (3)</td>
<td>Adults (18-60) (C)</td>
<td>1</td>
<td>*</td>
<td>2</td>
<td>1</td>
<td>4 (17%)</td>
</tr>
<tr>
<td>Health Promotion (4)</td>
<td>Adults (&gt;60) (D)</td>
<td>1</td>
<td>2</td>
<td>*</td>
<td>*</td>
<td>2 (8%)</td>
</tr>
<tr>
<td>None (E)</td>
<td>None (E)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>24 (100%)</td>
</tr>
</tbody>
</table>

* Questions within these categories are not available at this time.
The same training processes and scoring formats/rating scales are used for all events.

- Physician examiner criteria are common between sites and administrations

The following rating scale is used to establish the OSCE and TPx cut score:

- **Unacceptable** as compared to a recent graduate from a Canadian medical school
- **Borderline Unacceptable** as compared...
- **Borderline Acceptable** as compared...
- **Acceptable** as compared...
- **Above the level expected** of a recent graduate from a Canadian medical school
NAC Scoring and Reporting

- IMGs receive a Statement of Results from the NAC coordinating committee.
  - The report includes MCC branding and has the NAC logo.
- Provincial programs also receive the test results in more detail.
  - Some regions use total scores to rank order their applicants
  - Others use a pass or fail status – or both!
- Most events use their own registration systems and work with us to do administration.
  - MCC provides scoring and reporting to partners.
- The use of other requirements for selection into residency is at the discretion of the provincial programs.
  - Score reports describe only the NAC results.
Total score is a weighted combination of the OSCE and TPx components and is reported on a 100-point scale with a fixed pass score of 65.
Supplemental Feedback Report for Candidates in 2011

All candidates receive a Supplemental Feedback Report

For feedback purposes, the results from the examination have been grouped into ten competency areas: "History Taking", "Physical Examination", "Organization Skills", "Communication Skills", "Language Fluency", "Differential Diagnosis", "Data Interpretation", "Investigations", "Management", and "Therapeutics".

The results for these competencies were constructed from relevant ratings obtained across the 12 stations of the standardized candidate examination. Figure 1 displays low, borderline, and high levels of performance in each competency area. In addition, your obtained score on a particular competency is reported via a solid line within each shaded band. In turn, represents the standard variation in your scores if you were to be tested with other sets of stations covering the same or similar competencies. This information can be useful to you in order to assess your relative strengths and weaknesses across different competency areas.

To interpret this profile, look at where your performance band for a competency falls in relation to the borderline area. Most of your performance bands fall within or to the left of the borderline area, you know that scores in these areas are low. Most of your performance bands fall to the right of the borderline area, you know that your scores were higher in these areas. Small differences in the position of the performance bands or overlap between the bands in the score profile suggest that performance in these competency areas was relatively similar.

Figure 1. NAC Score Profile

*Please note that this approach to interpretation was tested on the candidates, it was not expected otherwise other than to make sure they understand the performance bands and therefore will be included in the supplemental report.
The NAC program is intended to allow for mobility between the jurisdictions.

- Centrally coordinated, regionally delivered.
- Common minimum eligibility requirements and use for selection into residency.

Residency directors also find the Supplemental Feedback Report useful. Currently, this is distributed at the discretion of each candidate.

Other criteria for selection into residency are up to each jurisdiction (e.g. educational report, clinical experience).

We advocate using the pass/fail status and urge programs to consider other score information with caution.
Cautions on Using Total Scores

- Understand what reported scores are NOT
  - % correct of all possible points on the test
  - Adjusted to enable a proportion of candidates to pass (some might call this ‘grading on a curve’)
  - Perfectly reliable
    - The same candidate, with the same preparation would obtain a different score when tested on an equivalent form.

- Possible error in scores depends on a variety of factors
  - Changes in the candidates performance on a given day
  - Quality of the presentation of the question, or clarity of response options
  - The number of times a skill or knowledge element is tested
  - The range of content domains covered in a single test
  - The match between the ability of the candidate and the difficulty of the content.
Thank-you for this opportunity to share some ideas.

Questions, Comments, Concerns?
# International Medical Graduates – the issue (2009 -1700 immigrant MDs - CIC)

<table>
<thead>
<tr>
<th>Country</th>
<th>Number</th>
<th>Country</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saudi Arabia</td>
<td>96</td>
<td>South Africa</td>
<td>82</td>
</tr>
<tr>
<td>India</td>
<td>72</td>
<td>Iran</td>
<td>63</td>
</tr>
<tr>
<td>Egypt</td>
<td>69</td>
<td>India</td>
<td>59</td>
</tr>
<tr>
<td>Iran</td>
<td>52</td>
<td>UK</td>
<td>37</td>
</tr>
<tr>
<td>UK</td>
<td>47</td>
<td>France</td>
<td>33</td>
</tr>
<tr>
<td>USA</td>
<td>42</td>
<td>Saudi Arabia</td>
<td>32</td>
</tr>
<tr>
<td>UAE</td>
<td>38</td>
<td>Pakistan</td>
<td>31</td>
</tr>
<tr>
<td>Kuwait</td>
<td>34</td>
<td>UAE</td>
<td>25</td>
</tr>
<tr>
<td>China</td>
<td>32</td>
<td>Sri Lanka</td>
<td>23</td>
</tr>
<tr>
<td>South Africa</td>
<td>31</td>
<td>Philippines</td>
<td>23</td>
</tr>
</tbody>
</table>

+++ 3000 or more Canadians studying abroad (potential 700/year)
Self-assessment (readiness?)
- MCC SAEE
- Saskatchewan Self-Assessment Study Guide

Examinations
- Evaluating Examination
- Qualifying Examination Part I
- Qualifying Examination Part II

Exam Preparation

Examination

Practice

Content & preparation
- Objectives
- $C^2$LEO
- Communication and Cultural Competence Program (CCC)

Results
- MCC Online
- Repository (PCRC)

Apply Online

Life-time access to repository

NAC OSCE!
MCC Evaluating Exam

REVISED IN 2008
FULLY COMPUTERIZED
CDM type of item

- All content is based on the MCC Objectives
  - Common clinical presentations
  - Clinical Task (diagnosis, management etc)
  - Patient age
  - ICD-10.
- CDM Items are authored using a ‘key features’ method
  - Tasks and options represent the elements most likely to lead to medical errors
  - Only the essential steps are included
  - Targets candidates general level of problem solving skills.
Converting Raw Scores to ‘Policy’ Passing Score

Example:
Pass on the raw score corresponds to a reported score of 65.
Some Final Cautions

- Different kinds of tests and items have their own error variances and contribute proportionally.
  - The MCQs are pretty reliable, the CDM may be less so!
- Total scores are most accurate at the passing score (by design).
  - Feedback on disciplines has low reliability.
- Estimates of errors increases in the lowest and highest score ranges.

- Use the pass/fail status!
- A typical relationship between accuracy and error in reported scores looks like this:

The recommended passing score may be here!
Cautions...

- It is important to keep in mind the exam is designed to be accurate for its specific purpose.
- Also, scores are transformed to the reported scale (linearly).
  - The error in the original scale may be magnified when the transformation is calculated.
Structure and Content of MCQ Modules (a.k.a. sections)

7 modules
28 MCQs

- Medicine: 4 questions
- Surgery: 4 questions
- OBGYN: 4 questions
- Pediatrics: 4 questions
- Psychiatry: 4 questions

PHELO
1 – Pop Health
3 - CLEO
4 questions and 4 pilots across each module.

Family Medicine crosses disciplines.
Disciplines have a mix of clinical tasks.