GUIDELINES for the DEVELOPMENT of OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE) CASES

November 2013
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td><strong>Step 1</strong> Reflect on the Purpose of the Exam</td>
<td>6</td>
</tr>
<tr>
<td><strong>Step 2</strong> Decide What You Want to Test</td>
<td>7</td>
</tr>
<tr>
<td><strong>Step 3</strong> Develop the Candidate Instructions</td>
<td>9</td>
</tr>
<tr>
<td><strong>Step 4</strong> Develop the Scoring Instruments</td>
<td>11</td>
</tr>
<tr>
<td><strong>Step 5</strong> Develop Case-Related Questions for Patient Encounter Probes</td>
<td>19</td>
</tr>
<tr>
<td><strong>Step 6</strong> Compile Information for the Standardized Patient (SP) and Trainers</td>
<td>21</td>
</tr>
<tr>
<td><strong>Step 7</strong> Describe Room Set-Up and Props</td>
<td>25</td>
</tr>
<tr>
<td><strong>Step 8</strong> Review, Revise and Pilot</td>
<td>26</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>27</td>
</tr>
<tr>
<td>References</td>
<td>28</td>
</tr>
<tr>
<td>Appendices</td>
<td></td>
</tr>
<tr>
<td>a. Checklist Example</td>
<td>29</td>
</tr>
<tr>
<td>b. Examples of Rating Scales Items</td>
<td></td>
</tr>
<tr>
<td>i. MCCQE Part II</td>
<td>30</td>
</tr>
<tr>
<td>ii. NAC Examiner Competency Descriptors</td>
<td>31</td>
</tr>
<tr>
<td>c. Rating Scale Resources</td>
<td></td>
</tr>
<tr>
<td>i. NAC Examiner Rating Guidelines</td>
<td>32</td>
</tr>
<tr>
<td>ii. NAC Examiner Rating Guidelines Example</td>
<td>33</td>
</tr>
<tr>
<td>d. Patient Encounter Probe (PEP) Example</td>
<td>34</td>
</tr>
<tr>
<td>e. General SP Responses in Role</td>
<td>35</td>
</tr>
<tr>
<td>f. OSCE Case Template</td>
<td>36</td>
</tr>
</tbody>
</table>
Dear reader-user,

This booklet is designed to provide guidance on the development and/or review of an Objective Structured Clinical Examination (OSCE) case for use in the Medical Council of Canada (MCC) performance-based examinations. Our hope is that this booklet will better prepare our content experts to develop cases that lead to a fair and valid assessment of candidates’ abilities.

Although case-writing is an iterative process which requires several rounds of review and revision, we have divided the process of OSCE case development into eight steps intended to be performed sequentially.

This booklet is not intended to provide a comprehensive review of all issues related to the development and administration of an OSCE. As such, important topics such as standard-setting procedures and rater training are beyond the scope of this publication.

Debra Pugh, MD FRCPC MHPE
Vice-Chair, MCC Central Examination Committee

Sydney Smee, PhD
Manager, Strategic Initiatives
Evaluation Bureau

November 2013
Background

An Objective Structured Clinical Examination (OSCE) is a performance-based test which allows for the standardized assessment of clinical skills. Since the OSCE format was first described by Harden in 1975 (1), it has been extensively studied and widely adopted by educational institutions and high-stakes testing organizations, including the Medical Council of Canada (MCC) (2-4).

During an OSCE, candidates are expected to perform a variety of clinical tasks in a simulated setting while being assessed by examiners using standardized rating instruments. Typically, candidates rotate through a number of stations in which they are expected to interact with a standardized patient (SP). Mannequins and simulation models are also used in OSCEs, although, to date, their use, or hybrid stations combining SPs with models, has not been incorporated into OSCEs at the MCC.

For each interaction, candidates are asked to demonstrate skills related to completing a physical examination, obtaining a history, interpreting data, or managing an emergent medical issue. A station may also assess aspects of effective communication or the ability to demonstrate professionalism. Several raters (usually one per station) are used to assess candidates’ performance, using station-specific checklists and/or rating scales.

Some of the advantages of this format are that it allows for:

• Direct observation of clinical skills
• Assessment of a broad range of skills in a relatively short period of time
• A fairer assessment based on a standardized approach
• Minimization of rater bias through use of multiple examiners

Considering that the development and administration of an OSCE is a laborious and resource-intensive endeavor, its use should be reserved for skills that cannot readily be assessed in a valid and reliable way using other examination formats.
OSCEs at the MCC

The MCC offers two performance-based examinations: the MCC Qualifying Examination (MCCQE) Part II (mcc.ca/examinations/mccqe-part-ii/) and the National Assessment Collaboration (NAC) examination (mcc.ca/examinations/nac-overview/). Both the NAC examination and the MCCQE Part II incorporate SPs in most of their stations. Some stations may also include a nurse or an allied health professional playing an assigned role. Both examinations use physician examiners to assess candidate performance with prespecified scoring instruments.

The MCCQE Part II was introduced in 1993 and is used to assess candidates’ competence with regards to the knowledge, skills and attitudes essential for medical licensure in Canada prior to entry into independent clinical practice. The exam content is based on common and/or critically important patient presentations. This OSCE currently consists of a 12-station series comprised of eight ten-minute and four coupled five-minute (couplet) stations. In the couplet stations, candidates have five minutes for a clinical encounter and five minutes to either prepare for the clinical encounter or to answer written questions related to that encounter. The non-patient component of these stations is referred to as a Patient Encounter Probe, or PEP.

The National Assessment Collaboration (NAC) was formed as “an alliance of Canadian organizations streamlining the evaluation process for international medical graduates (IMGs) seeking a license to practice medicine in Canada.” As its first initiative, the NAC developed an examination to assess IMGs readiness for entry into a Canadian residency program. The resulting OSCE was first administered in 2010. In its current state, the NAC examination consists of 12 stations (of 11 minutes each) based on clinical scenarios. The clinical scenarios sample from problems in Medicine, Pediatrics, Obstetrics and Gynecology, Psychiatry and Surgery. Performance on each scenario is assessed on up to nine possible competencies, including: history-taking, physical examination, organizational skills, communication skills, language fluency, diagnosis, data interpretation, investigations, management, and therapeutics.

Purpose

The purpose of this booklet is to provide authors with guidance on the development of cases for either of the performance-based examinations administered by the MCC. Our hope is that this booklet will guide authors in the creation of high-quality OSCE cases for the fair and valid assessment of candidates’ abilities.

The contribution of content experts as case authors is invaluable to the work we do at the MCC. As engaged clinicians and educators, our authors offer not only their expertise with regards to content, but also a dedication to excellence in assessment.

To this end, the process of OSCE case development has been deconstructed into eight steps, and we advocate that case writers follow these steps sequentially. Following these steps will help to ensure that issues related to content, feasibility, authenticity and reproducibility are addressed early in the development of cases.
Before beginning to develop an item for any exam, including a case for an OSCE, one must consider the purpose of the exam. This includes being clear on whether the exam will be formative (i.e., an interim examination used to provide feedback to learners) or summative (i.e., an end of rotation examination used to assess competence).

The expected ability or clinical performance level of the candidates must be considered to ensure that the item content and the task difficulty level are appropriate. The MCC administers two high-stakes summative OSCEs. Both are administered nationally, at multiple sites across the country, each with its own purpose:

- **MCC Qualifying Examination (MCCQE) Part II**: Assesses the candidate’s core abilities to apply medical knowledge, demonstrate clinical skills, develop investigational and therapeutic clinical plans, as well as demonstrate professional behaviors and attitudes at a level expected of a physician in independent practice in Canada.

- **National Assessment Collaboration (NAC) Examination**: Assesses the readiness of an international medical graduate (IMG) to enter into a Canadian residency program.
STEP 2

DETERMINE WHAT YOU WANT TO TEST

Having reflected on the purpose of the exam and the level of clinical skills being assessed, the second step for the case author is to clearly define the construct of interest. One must consider the assessment objectives for the OSCE test item or case that they are about to write. For example, one might choose to develop a case that will assess “the ability of a candidate to generate an appropriate differential diagnosis when examining a patient presenting with an acute abdomen.”

Not everything could or should be assessed using an OSCE. OSCEs are resource-intensive, and there are more efficient means of assessing more knowledge-based and clinical reasoning skills (e.g., using multiple-choice or short-answer questions). An OSCE is best suited to assessing skills that are not easily measured with other methods. Commonly, OSCEs are used to assess history-gathering, physical examination, communication, interpretation of data, and/or management skills. In addition, OSCEs can and have been used to assess all the CanMEDS roles (5-7).

The content sampling specifications, or blueprint, of an exam should be carefully developed to ensure that the sample cases for any one test form represent the broad scope of content and domains to be assessed (8). Generally speaking for MCC OSCEs, the blueprint specifies that a test form should sample from all the major clinical disciplines (i.e., Pediatrics, Medicine, Surgery, Obstetrics and Gynecology, and Psychiatry). Content may also be related to Population Health as well as the Ethical, Legal, and Organizational aspects of medicine.

While the MCC’s blueprint specifications are currently undergoing a major review, the underlying principle is that patient problems for any one test form should sample broadly across the spectrum of common and important clinical presentations, across patient demographics, and across clinical skills. This
underlying principle is being reframed and will likely change the emphasis on what is assessed in a given test form, but it will not disappear from the specifications. In other words, the range of patient problems and clinical tasks that are appropriate for an author to develop is, and will remain, broad.

In addition, a variety of tasks should be included in each test form and, depending on changes that may occur to the format of either examination over time, may be combined within a station. Pertinent tasks include history-taking, physical examination, communication, and management skills. A test form should also sample as broadly as possible across presenting problems and/or body systems with consideration for patient demographics, such as gender and age groups. For example, one would ideally not include multiple cases related to diabetes mellitus in one test form, given that the total number of cases is usually only about 10 to 15.

The MCC uses the Objectives for the Qualifying Examination to define the clinical scope of the domain being sampled by Part II. The Objectives also serve as a general overview of possible test material that may appear on the NAC Examination (apps.mcc.ca/Objectives_Online). The Objectives have been defined in behavioral terms, and reflect the MCC’s expectations of competent physicians. They deal with data gathering, diagnostic clinical problem-solving, and the principles of management, which are applicable, in whole or in part, to the range of clinical situations commonly faced by physicians.

Ideally, an author will use a real patient as an inspiration for writing a case. Authors who are working from a real case are more likely to include relevant and consistent details for a specific presentation and to avoid the tendency toward writing a generic textbook case. Realistically, authors more often work from a composite of patient cases, which still serves the purpose of supporting the creation of a specific presentation of a problem. Although completely fictitious cases can and have been used, in our experience, this tends to be the least successful approach to producing a workable OSCE station.

Props are often used in OSCE cases. For example, a candidate may be asked to obtain a history from a patient presenting with abdominal pain, then review an abdominal radiograph related to the case, or they may be asked to interpret an electrocardiogram as they manage an acutely ill patient. If any props such as radiographs or electrocardiograms are required, they should be obtained before the case is developed, as trying to find appropriate images after-the-fact is often frustrating. Inevitably, a case will be dropped for a lack of an appropriate image or will have to be re-written, sometimes drastically, to match the image that is found.
The candidate instructions provide the information to candidates prior to entering an OSCE station and include the clinical stem and task. The candidate instructions, or “stem,” should include the setting, the patient’s name, gender and age, as well as their presenting complaint or reason for referral. It should also include the “clinical task” that must be performed by the candidate, as well as the time allotted for the task.

The stem, including the required clinical task, should be written in a clear and unambiguous manner. “Explore this further with the patient” is an example of an ambiguous phrase that could be more clearly stated as “Take a focused history and address this patient’s concerns.”

The language should be appropriate and geared to the individual who is providing the information. For example, a patient who is reporting a symptom will likely use phrases such as “I feel short of breath” or “I can’t catch my breath,” whereas a patient referred by a colleague may be described as having “dyspnea.” Similarly, a patient may describe their skin as “yellowish,” whereas a clinician would be more likely to use the term “jaundiced.”

Other pertinent information, such as features on history, physical examination, or investigations, should be included when and if appropriate. The amount of information provided a priori will depend on the required task and the candidate’s level of ability. For example, if the case objective is to assess the candidate’s ability to perform a focused physical examination of a patient with chronic liver disease, it would be appropriate to provide the candidate with a synopsis of the history as well as vital signs:
Jennifer Hawkins, a 46-year-old woman with a history of hepatitis C and chronic alcohol abuse, has been admitted to hospital for investigation of elevated liver enzymes. Her vital signs are: heart rate 90/min, blood pressure 108/84 mmHg, respiratory rate 12/min, oxygen saturation 99% on room air, temperature 37.2°C.

In the next 10 minutes, perform a focused physical examination to look for signs of chronic liver disease.

However, if the objective is to assess a candidate’s ability to assess and manage a patient presenting with confusion related to hepatic encephalopathy, then it may be more appropriate to provide minimal information in the stem:

Jennifer Hawkins, a 46-year-old woman with a history of cirrhosis, presents to the emergency department with a three-day history of confusion.

In the next 10 minutes, assess and manage this patient.

The wording of the “clinical task” can also affect the case difficulty. For example, “perform a focused physical examination to look for signs of chronic liver disease” is a more specific and thus simpler task than “assess and manage the patient,” which requires the synthesis of information as well as clinical judgment.

The amount of time available for the task must also be taken into account when developing a case. The MCCQE Part II has both five and 10-minute stations, while the NAC examination has 11-minute stations. It might not be reasonable to expect a candidate to perform an appropriate history and physical examination of a patient presenting with chronic fatigue in a five-minute station; however, a knee exam can easily be performed in that time frame. The complexity of the case should be adjusted to reflect the time constraints so as to allow for a fair assessment of a candidate’s ability.

Other practical considerations when developing a case include ensuring the case is as plausible and realistic to the candidate as possible. For example, it is more realistic to frame the case so that it is the candidate’s first time assessing the patient (i.e., in the emergency department or in a walk-in-clinic, or for a colleague who is absent) rather than asking them to pretend that they have a pre-existing relationship with the patient. Because OSCE cases may be used across different time zones and at different times of the year, it is also important to ensure the case works in real time (e.g., by stating that the symptoms started “three hours ago” rather than at “9 a.m.”, or “two months ago” rather than “in January”).
STEP 4
DEVELOP THE SCORING INSTRUMENTS

Once the candidate instructions have been developed, the instruments that will be used to score performance should be developed. Scoring instruments may consist of a checklist or rating scale/s or both. For the MCCQE Part II, a combination of checklists and rating scale items is common. The NAC examination relies solely on a preset series of rating scales associated with competencies (e.g., history-taking, physical examination, communication skills), many of which apply to each case and are supported by case-specific performance guidelines. For each case, the author can select up to ten of these rating items for scoring.

Checklists – MCCQE Part II

Checklists are useful in the assessment of clinical skills and provide an objective and standardized tool that requires little examiner training to complete. In order to be useful, however, checklists must be carefully constructed and tailored to the required clinical task (see Table 1).

Number of Checklist Items

The number of items required on a checklist depends on the case and the time allotted. Generally speaking, for short stations (five to seven minutes), eight to about 25 checklist items are acceptable. More items may be appropriate for longer stations or for stations where there are many steps, as may occur in some management stations where ordering a number of tests would require multiple items (e.g., order a CBC, order electrolytes, order an ECG, order a chest radiograph). Each item should be tailored to the clinical context. For example, there may be no items, one item, or several items regarding alcohol consumption, depending on its relevance to the presenting problem and assigned task. Regardless of the number, the examiner should only include items that are relevant to the assessment of candidate’s ability. A checklist should not be exhaustive.
Checklist Compatibility with Required Task

As described above, the candidate instructions outline the task required of candidates. The checklist should reflect this closely. For example, if candidates are instructed to examine a patient with abdominal pain, the checklist should be composed of physical examination items, not history-based items. If the task requires candidates to counsel a patient about smoking cessation, the checklist items should be related to key elements of the counseling task rather than information-gathering.

Checklist Item Characteristics

Checklists are most useful when assessing thoroughness and/or key elements of a task. (See Appendix a for an example of a checklist.) When developing a checklist, one must ensure that it is comprised of items that are discrete, observable and dichotomous.

Discrete

Each item should represent only one concept. For example, an item should not include two symptoms such as “enquires about frequency and severity of anxiety symptoms.” In this example, it should be divided in two separate items. If it is deemed necessary to combine several points in one item, specific instructions should be provided to examiners with regards to scoring. For example, “Bubble in only if the candidate asks about at least three of five symptoms.”

Observable

Examiners should only be asked to make judgments on observable skills; e.g., “palpates axillae for lymphadenopathy.” In contrast, it would be difficult for examiners to assess candidates on items that call for judgment such as “understands the impact of alcohol abuse on the patient’s life,” or “appreciates the severity of the patient’s diabetes.”

Dichotomous

The assessed items in the checklist are currently scored dichotomously (i.e., they were either performed satisfactorily, or they were not completed satisfactorily). Examples of dichotomous checklist items include: “assesses vibratory sensation in the lower extremities,” “enquires about previous blood transfusions,” or “identifies ST-elevation on electrocardiogram.” The use of checklists with “attempted” or “done satisfactorily” items have been piloted with encouraging results, and have just been implemented for some cases. If items in your checklist seem suitable for scoring as “attempted” or “done satisfactorily”, then provide that direction. Items related to continuous variables, such as establishing rapport or demonstrating professionalism, are better assessed using a rating scale.
Checklist Language

Each checklist item should begin with an action verb (e.g., “enquires about visual hallucinations,” rather than “visual hallucinations”) to better guide examiners in terms of what is expected of candidates. If necessary, clarification can be provided. For example, if “asks about gynecological history” is a checklist item, you may wish to be explicit about what is expected:

- Asks about age at menarche
- Asks about date of last menstrual period

Table 1: Checklist Considerations by Station Type

<table>
<thead>
<tr>
<th>STATION TYPE</th>
<th>CONSIDERATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Examination</td>
<td>Items may direct the examiner to report a finding to candidates (e.g., “BP is normal.”) if initiating the action is sufficient for a candidate to receive credit.</td>
</tr>
<tr>
<td>History</td>
<td>Verbs matter and should guide the examiner. For example, elicits indicates some latitude on how a candidate gathers the information from the patient; asks indicates that a candidate must be specific; e.g., “Asks about use of Aspirin®” versus “Elicits medication history.”</td>
</tr>
<tr>
<td>Communication (e.g., counseling or patient education)</td>
<td>Items may include information to be gathered from the patient and must include items indicating the education and/or advice, and/or support to be given to the patient. SP instructions may include questions to ask all candidates (i.e., standardized cues to specific challenges to assess all candidates’ ability to respond).</td>
</tr>
<tr>
<td>Management</td>
<td>May be acute care (e.g., ED setting) or may require decisions about patient management (e.g., solving problems related to multiple medications). The checklist may include items on history, physical examination and/or communication, but must include items regarding decisions, orders, treatment initiatives, etc.</td>
</tr>
</tbody>
</table>

No Killer Items

Candidates sometimes perform what are considered dangerous or egregious acts in the course of a performance-based examination. This type of response can be difficult to assess in a standardized way. One approach to help identify grossly incompetent candidates is to build “killer items” into cases (i.e., if a candidate commits a fatal error, they score zero on a station even if they have satisfactorily completed other items). Examples of potential killer items include the failure
to recognize that a patient is in shock or the administration of Insulin in a patient with hypoglycemia. The ramifications of these items within an OSCE are significant, and killer items are not currently used by the MCC. Instead, examiners are expected to flag and describe any such actions in a designated space on all mark sheets. A flag, for dangerous or professionally inappropriate actions by any candidate who passes, triggers a focused review of their examination results and the potential reconsideration of their pass result.

Oral Questions

In any station, a case author may include an oral question (or more than one) that examiners must ask all candidates to answer during the last minute(s) of the station. Alternatively, the SP or the nurse (if present) may ask the oral question during the encounter.

Adding oral questions is feasible when the clinical task can be accomplished in less time than required for the station and when questions that further assess clinical competence relative to the specific presentation can be asked; e.g., asking candidates for their working diagnosis or an initial treatment plan. The question should be directly related to the patient scenario and may relate to problem-solving or data synthesis (e.g., establishing a diagnosis or management plan), or to non-clinical skills (e.g., truth-telling).

The most challenging aspect of including an oral question is developing an answer key that is likely to capture all acceptable correct answers, as it requires authors to anticipate candidates’ responses. Oral questions work best when there are a limited number of possible correct responses.

Generally speaking, in the MCCQE Part II checklist, the correct answer is provided first, followed by other anticipated options. These other options may be scored as zero, but will aid examiners in recording how the candidate responded. The last option is always “Other answer or no answer given.”

**Example:**

“What is the most likely diagnosis?”

- Small bowel obstruction (1)
- Perforated viscus (0)
- Cholecystitis (0)
- Other answer or no answer given (0)

However, in the case of making a diagnosis, specifying only one answer may not be possible if the case is lacking sufficient information to be that specific. One option is to have the examiner provide more information.
**Example:** “Initial investigations reveal a TSH of < 0.01 mIU/L and a free T4 of 28.99 pmol/L. What would be your next step in management?”

- Start PTU or Methimazole (1)
- Start Propranolol (0)
- Order radiiodine (131-I) (0)
- Administer Lugol’s solution (0)
- Other answer or no answer given (0)

Alternatively, there might be several equally correct responses, and this can be reflected in both the question and the answer key.

**Example:** “What is your differential diagnosis? Give up to three answers.”

- Major depressive disorder (1)
- Bipolar disorder (1)
- Depression not otherwise specified (1)
- Other answer or no answer given (0)

Oral questions can also be used to ask legal or ethical questions related to the patient.

**Example:** “The patient’s employer calls your office to request information about the patient’s illness, as he suspects she is malingering. How will you address his concerns?”

- Inform the employer that you cannot comment on his employee’s health (1)
- Reassure him that you do not believe the patient is malingering (0)
- Agree with him that the patient is malingering (0)
- Other answer or no answer given (0)

A well-constructed oral question will elicit higher order skills requiring the synthesis of clinical information gathered from the patient. If this is true, increasing the item weight may be appropriate in order to reflect the higher order of skill required. In the NAC examination, only correct answers are specified but they are not weighted. Instead, the examiner includes consideration of the given answer(s) when rating items on the rating scale.

**Key Feature Approach to Checklists**

As noted above, a common criticism of checklists is that they tend to reward thoroughness over expertise (9). Because of this, it is theoretically possible for a less competent candidate, using a rote or shotgun approach in the OSCE, to outperform a candidate with more knowledge and expertise who tailors their approach to the clinical problem. One possible solution to this issue is to develop
items based on key features. Key features focus on: (1) critical or essential steps in the resolution of a problem; (2) steps in which candidates are most likely to make errors in the resolution of the problem; or (3) a difficult or challenging aspect in the identification and management of the problem in practice (10-13).

Using this method, a checklist can be refined to include only those elements that are likely to discriminate between competent and non-competent test takers. For example, in the case of a patient with back pain, a case writer may create key feature items that recognize identifying symptoms suggestive of cauda equina and not have items related to the duration and location of the pain.

The use of key feature items may only significantly decrease the number of items on a checklist (i.e., to less than ten), which can be disconcerting to examiners accustomed to traditional checklists. However, if balanced by the inclusion of relevant behavioural rating scale items and focused examiner orientation, then examiners will hopefully make the transition. The key feature approach is becoming the preferred method for creating new test items at the MCC, and authors are encouraged to use this framework in creating checklists.

**Checklist Item Weighting**

Previously, MCCQE Part II checklist items were weighted based on their relative importance according to expert judgment. This approach was used to compensate for the tendency of checklists to reward thoroughness over expertise, as more important items were worth more points. In such longer lists with no weighting, the “signal” from important and discriminating items was drowned out by the “noise” from many less important items. In this manner, weighting items required substantially more effort (i.e., to achieve consensus on the relative weights and to program scoring applications) and, while weights may have contributed to score validity, the weights did not contribute to score reliability.

As of 2013, in almost all instances, checklist items for the MCCQE Part II are scored as one or zero, as this is the simplest approach for development and aligns the scoring instruments with automated scoring processes. Although shorter checklists are a result of the current key feature approach, consideration must be given toward balancing checklist items and rating scale items relative to the clinical context. See below for further discussion of this topic.

**Rating Scales**

In contrast to checklists, rating scales are useful in the OSCEs for assessing behaviors along a continuum (e.g., organization of approach) or for rating the ability on a task as opposed to the details of how it was done (e.g., using a
rating scale item for a history-taking skill vs. a checklist of questions). Rating scales allow experts to use their judgment in rating various elements of a task, and thus are ideal for more complex tasks and abilities such as organizational skills, establishing rapport, and attentiveness to verbal and non-verbal cues. More holistic judgments may also be fairer for candidates in any one cohort who may have a wide range of experience and expertise. For these reasons, the NAC examination relies on rating scales to assess specific competencies, such as history-taking, physical examination, investigations, data interpretation, and management.

Because rating scales require examiners to use considerable judgment, it is important to minimize the inherent subjectivity of rating as much as possible. Some ways to do this include: providing clear instructions, ensuring adequate rater training, and providing behavioral anchors and other supporting guidelines. Behavioral anchors are descriptive examples of the range of performance for each item and can improve inter-rater reliability (14). Rating scales also require more time to score and are best scored at the end of an encounter. As the time between OSCE stations is often short, the number of rating scale items should be limited – one or two to a maximum of seven in the MCCQE Part II and from seven to nine for the NAC examination.

Rating Scales – MCCQE Part II

The number of rating scale items used in each MCCQE Part II case will vary. Case authors select the appropriate rating scale items from a preset bank. For example, in a case with a strong physical examination component, the candidate might be assessed almost completely using a checklist, with only one or two rating scale items (e.g., organization of physical examination and demonstration of technical skills). In a case focused on communication skills, however, the candidate might be assessed using several rating scale items designed to assess rapport with patient, questioning skills, professional behavior, etc., and only a few key feature checklist items. (See Appendix b(i) for examples of the behaviorally-anchored rating scale items currently available for MCCQE Part II cases.) Since rating scale items have a maximum score of five and checklist items are worth one score each, case writers need to weigh the purpose of the station with the relative importance of the checklist items and the rating scale competencies for each case. For instance, with five key feature checklist items and three rating scale items, the checklist items would represent only 25% of the total score possible for that station, while the rating scale would represent 75% of the total score (see example below). This may or may not represent the clinical importance of the key features relative to the rating scale item competencies.
Rating Scales – NAC Examination

Like the MCCQE Part II, the NAC examination relies upon a preset series of rating scale items to assess critical competencies. For each case, the author needs to select up to nine of the following competencies for scoring: history-taking, physical examination, organizational skills, communication skills, language fluency, diagnosis, data interpretation, investigations, management/therapeutics. Candidates’ performance is rated relative to what is expected of a recent graduate of a Canadian medical school, using the competency descriptors (see Appendix b(ii)).

In addition to the item anchors, there are case-specific guidelines which provide an explicit (but not exhaustive) list of clinical skills expectations at a particular performance level. These guidelines are meant to help examiners discriminate between five levels of performance (ranging from “unacceptable” to “above level of a recent graduate”). A case author will provide additional case-specific guidelines to assist physician examiners in making valid and reliable judgments, as the case-specific guidelines are intended to keep examiners aligned with the case’s intended purpose. See Appendices ci and cii for further details regarding examiner guidelines for the NAC examination.
**DEVELOP CASE-RELATED CONTENT and QUESTIONS for PATIENT ENCOUNTER PROBE (PEP) (if applicable)**

Just as oral questions can be incorporated into an OSCE case and scored in the checklist, questions can be asked after the SP encounter. The MCCQE Part II uses four such stations, coupling a patient encounter with tasks to be completed beforehand or afterwards (e.g., admission orders or a written task). These are referred to as Patient Encounter Probes (PEPs). The purpose of the PEPs, as with the oral questions, is to further assess clinical competence relative to the specific presentation.

**Example of PEP questions following the patient encounter**

**Q1** A paracentesis reveals a serum-to-ascites albumin gradient of 12 g/L. Given this patient’s history, what is the most likely etiology?

<table>
<thead>
<tr>
<th>Score</th>
<th>Cirrhosis</th>
<th>Alcohol-related liver disease or hepatitis</th>
<th>Heart failure/CHF/pericarditis</th>
<th>Liver metastases/cancer/hepatocellular carcinoma</th>
<th>Budd-Chiari</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Maximum 4

**Q2** How would you best manage this patient’s symptoms in the long term?

<table>
<thead>
<tr>
<th>Score</th>
<th>Diuretics</th>
<th>Dietary salt restriction</th>
<th>Paracentesis</th>
<th>Fluid restriction</th>
<th>Liver transplant</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Maximum 5

See Appendix d for further PEP examples.
PEPs that require candidates to complete a task prior to seeing the patient are not scored directly. For example, the candidate may be asked to review a patient’s chart prior to meeting with the patient. In this example, the case author creates the patient chart. Ideally, the chart includes information that will assist the candidate in the assigned patient task. Other information not immediately relevant to the meeting with the patient should be added, as part of the candidate’s task is to be able to identify and use the relevant information within the set time frame.

The following are general principles to consider when creating a written component (i.e., PEP) for an OSCE station:

1. The questions must be inextricably linked to the specific patient problem and not be comprised of generic questions that can be tested in a multiple-choice examination. Remember that generic content can be far more effectively and less expensively assessed with other methods and should not be incorporated to an OSCE case.

2. The questions or task should be answerable within the allowed time frame.

3. Questions and props should not cue answers to other items within the same component; e.g., do findings on an x-ray inadvertently answer more than one PEP question? Similarly, subsequent questions should not cue the candidate to previous answers.

4. The answer key should be comprehensive, and should allow for standardized and consistent interpretation by different markers. It should also list all the likely answers (including those that are worth zero), and be explicit with regards to scoring expectations. For example, if the answer “colon cancer” is scored with two marks, does “cancer” get two marks, one mark or zero?
The SP information should be complete yet concise; it should be written in layperson terms based on the patient’s profile (e.g., educational level) to make it easier for SPs to use language tailored to the specific patient and to ensure that SP trainers and SPs understand the problem from the patient’s perspective.

**Demographic data**
Specify demographic data for those recruiting SPs. The age (best provided as a range, like 35-40 years old), gender and other specific characteristics (e.g., non-obese, no abdominal surgical scars) should be provided. Consideration should be given to recruitment challenges when distinguishing traits are required (e.g., a particular ethnicity).

**SP starting position**
Specify where the encounter is taking place: doctor’s office, emergency department, health clinic or other, and how the patient should be positioned at the start of the station (e.g., in the ED, lying on his back on a stretcher with his legs flexed).

**Appearance**
The SP should be provided with specific information on how to dress based on the presenting problem and the patient’s demographic. Should the patient’s appearance be professional, casual, or disheveled? Should the SP be wearing a hospital gown and have a sheet available? If the chest needs to be exposed, you may wish to instruct female SPs not to wear a bra. If the feet should be exposed, the SP should be instructed not to wear socks (unless it is part of the candidate’s task to expose the area). Specify if the SP is to have a “personal” prop such as a cane or a pregnant belly. If the SP requires make-up, this should also be specified (e.g., abrasion to chest, facial pallor and clamminess, or none). Standardizing the quality of make-up application across various sites is challenging, therefore the use of photographs and illustrations is encouraged for make-up application.
Behavior, affect and mannerisms

The SP should be provided with some direction on how to behave during the encounter. The more specific the instructions, the better. For example, “pacing around the room” may provide more guidance than simply stating “restless.”

The SP’s concerns and/or perception of their problem (e.g., serious, not serious), their expected response to the candidates (e.g., good eye contact, argumentative, attentive), and their general affect (e.g., tense and anxious) should also be included.

Pay particular attention to scenarios where the SP’s affect might change during the encounter. In order to standardize the patient portrayal, the SPs must have as clear an understanding as possible of the specific candidate responses that would cue their own responses or behavior, as well as timing cues (e.g., must say a given statement at five-minute mark). For instance, if the SP is meant to reveal certain information only if the candidate approaches the patient in a certain way, then state clearly what the candidate must say or do to get that result (e.g., the SP will only disclose that they are in an abusive relationship if the candidate is able to establish rapport with the patient by reassuring them of confidentiality). Specifying these details enables the SP to recognize when that condition has been met.

Opening statement

For many cases, the SP will be required to make a brief statement when the candidate first enters the room or addresses them. Basic, common language should be used (e.g., “I’m worried about this pain in my leg”), and shorter statements are usually best.

Questions the SP must ask

“Must ask questions” are prompts that the SP must ask all candidates. For example, “Should I keep taking my mom’s Tylenol 3® at night?” These questions provide a standardized cue to candidates in order to assess their specific knowledge and/or attitudes as well as their ability to respond to the issue at hand.

Ensure that you are specific about the timing of the “must ask” question; for instance, when the SP must ask the question at the final warning buzzer or earlier in the case (making sure the candidate has time to actually answer the question!), or upon a certain line of questioning (e.g., when the candidate asks about exacerbating factors, the SP must ask “Should I stop doing yoga?”). Not all cases will or should have “must ask” questions. If the SP is not given a specific, scripted question to ask, they will be trained to not ask any questions.
**Physical findings**
While the SPs generally have no abnormal physical findings, they can be trained to reliably portray some of them (but not all; e.g., cannot simulate a heart murmur). Examples of findings that can be simulated include: muscle weakness, joint tenderness, confusion, seizures, gait abnormalities, tremor. If the SPs are supposed to be in pain, they need information on how to respond appropriately; e.g., wincing when touched in a certain area or walking with a limp. If distracting maneuvers are anticipated (i.e., if a candidate is trying to differentiate between organic pain and malingering), the SP should be instructed on how to demonstrate the findings consistently. Descriptions of symptoms in terms of triggers from daily living and range of motion are important additions to physical examination maneuver responses and for consistent SP performance.

**History of presenting problem**
Information must be from the patient’s perspective, in the patient’s language, and include enough information so that the SPs can reliably and realistically answer a wide range of questions regarding their problem. For example, you should include information about: onset, duration, progression and frequency of symptoms; location, radiation, quality and intensity of pain; alleviating/aggravating factors; precipitating incident; and associated symptoms.

Since examinations are administered over the course of a day and in different seasons, it is important to choose general statements about place and time rather than being too specific. This ensures a more standardized response, thus reducing the chances for confusion. For example, rather than stating that the symptoms started “in July,” “Saturday” or “at breakfast time,” it is more appropriate to state that they started “three months ago,” “three days ago” or “three hours ago.” If pain is exertion-induced, it is easier to say “it gets worse when I do yard work” rather than “shoveling snow” or “working in the garden.”

**Relevant past medical history**
Focus should be on pertinent positive information related to the purpose of the case. An unnecessarily detailed history complicates training and makes it more difficult to standardize SP portrayals. Include (as appropriate) information about: past illnesses; current and past medications; allergies; hospitalizations; accidents/injuries; and/or other considerations. A simple timeline from onset to current moment is often helpful.

**Relevant social history**
As above, focus on pertinent information about the patient by attempting to find the balance between too much and too little information. Include
(as relevant) information about: living environment (where, with whom); sexual history; drug, alcohol and smoking habits; and/or other considerations relevant to the purpose of the case. Most of the time, it is not necessary to list specific information such as the spouse’s name or the name of the workplace/school. SPs will be trained to use familiar names they can easily remember, as there is no need for that kind of information to be standardized.

**Relevant family history**

Unless otherwise specified, SPs are instructed to give benign, “boring” family histories. **DO** provide any pertinent positive information about parents, siblings or other family, but only as necessary. Simple information is helpful; e.g., parents alive and well, or uncle died of a heart attack at age 46, or no family history of diabetes, or niece had febrile seizures as an infant.

**Critical review of systems**

Only provide information NOT included in the above sections and only provide information critical to a particular patient’s presentation. Generally, only pertinent positives are relevant. Two exceptions would be providing responses for an uncommon or less frequent pertinent negative and providing a response for a finding that the SP may be asked about frequently (e.g., no pain on swallowing; no memory impairment). Clarifying negative reactions in these instances reassures the SP that these issues were addressed in their training.

**Additional information to trainers and SPs**

Providing additional information such as common questions from candidates (e.g., level of exercise) and general SP responses can be very helpful. See **Appendix e** for guidelines to general SP responses for the NAC examination. As well, providing some less common but important questions (e.g., family history of blood clots), brief descriptions of different physical examinations (e.g., maneuvers to differentiate between joint and soft tissue injuries), and any tips you think would help with simulating the affect and symptoms will enable more consistent SP performances. Remember that if the information isn’t given to the SP in the case information, they will answer the question as benignly as possible (i.e., “No, I don’t have that symptom” or “It’s normal.”).

**Cross-reference checklist to SP instructions**

Do ensure that responses to all scoring items in the checklist are reflected in the SP script. Demonstrate the link by adding key words from the SP’s responses next to each item on the scoring instrument, even if it is simply “Normal” or “Negative.”
Since the case-writer will not be involved in the set-up or administration of the OSCE, it is important to provide specific instructions to staff about the room set-up, props and required equipment.

For example, specify whether an examining table is required (and whether it should be accessible from the right side of the patient, or whether the head of the bed should be raised), the number of chairs required (include one for the examiner), the props to be used (such as an electrocardiogram or a radiograph), and special equipment (such as a sphygmomanometer or a reflex hammer). Other considerations may include props that relate to the realism of the scenario, such as intravenous poles, urinary catheters, or nasal prongs.

Consider providing reference material for the trainers, SPs and/or examiners. This could include photographs, diagrams and figures, or journal articles (referenced, as appropriate).
Review, Revise and Pilot

Case development is an iterative process that requires careful thought, review and revision. Even after investing your time and effort in a case, it is important to remain open to feedback. Having input from clinical colleagues and staff can only improve the quality of a case and may identify any potential problems early in the process. For instance, role-playing the case with a colleague may allow authors to identify missing information from the SP script. And do note that even after all this review, a case may still be piloted in a live examination to determine how well it functions, both logistically and psychometrically.

Conclusion

We hope that this booklet has helped to deconstruct a potentially difficult and time-consuming task into meaningful steps, and that it will support you in the development of high-quality OSCE cases.

See Appendix f for an OSCE case template.
References

How to reference this document:


ACKNOWLEDGEMENTS

We would like to acknowledge the following individuals for their help in reviewing and editing this work: Dr. Joan Fraser, Ingrid de Vries, Alexa Fotheringham, Judy McCormick, and Dr. Claire Touchie. We would also like to thank all the staff at MCC who contributed to reviewing, editing, formatting and translating this document.


APPENDIX

Candidate’s Instructions:
Luc Léger, 59 years old, presents to your office complaining of jaundice. In the next 5 minutes, obtain a focused and relevant history. At the next station, you will be asked questions about this patient.

Examiner’s Checklist

Fill in the bubble for each item completed satisfactorily

- 1 Elicits onset/duration
- 2 Elicits progression
- 3 Elicits associated symptoms
- - dark urine
- - pain
- - color of stool
- - fever
- 4 Elicits risk factors
- - previous exposure to hepatitis
- - recent blood transfusion
- - intravenous drug use
- - foreign travel
- 5 Elicits a history of alcohol use
- 6 Conducts a review of systems
- - skin
- - gastrointestinal
- - weight loss
- - change in appetite

Did this candidate demonstrate a lapse in professional behavior?

No ____ Yes ____ (If yes, why?) ________________________________________________
- Disrespectful to others (e.g., to patient, nurse)
- Over-investigated / over-managed the patient
- Actions raised ethical and / or legal concern

Briefly describe the behavior for any of the above reasons or any other observed lapse: __________________________________________________________________________
## APPENDIX

### Examples of Rating Scales Items

**MCCQE Part II**

#### Listening skills

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interrupts inappropriately, ignores patient’s answers</strong></td>
<td>Impatient</td>
<td>Borderline unsatisfactory</td>
<td>Borderline satisfactory</td>
<td>Attentive to patient’s answers</td>
<td>Consistently attentive to answers &amp; concerns</td>
</tr>
</tbody>
</table>

#### Questioning skills

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Awkward, exclusive use of leading or closed ended questions, jargon</strong></td>
<td>Somewhat awkward, inappropriate terms, minimal use of open-ended questions</td>
<td>Borderline unsatisfactory, moderately at ease, appropriate language uses different types of questions</td>
<td>Borderline satisfactory, moderately at ease, appropriate language, uses different types of questions</td>
<td>At ease, clear questions, appropriate use of open and closed ended questions</td>
<td>Confident and skilful questioning</td>
</tr>
</tbody>
</table>

#### Organization of interview

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scattered, shot-gun approach</strong></td>
<td>Minimally organized</td>
<td>Borderline unsatisfactory, flow is somewhat logical</td>
<td>Borderline satisfactory, logical flow</td>
<td>Logical flow with sense of purpose</td>
<td>Purposeful, integrated handling of encounter</td>
</tr>
</tbody>
</table>

#### Organization of physical examination

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scattered, patient moved unnecessarily</strong></td>
<td>Minimally organized</td>
<td>Borderline unsatisfactory, flow is somewhat logical</td>
<td>Borderline satisfactory, logical flow</td>
<td>Logical flow with sense of purpose</td>
<td>Purposeful, integrated handling of examination</td>
</tr>
</tbody>
</table>

#### Demonstration of technical skills

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No skill</strong></td>
<td>Manoeuvres too rushed or clumsy, unlikely to provide reliable / useful information</td>
<td>Borderline unsatisfactory: Some skill but likelihood of reliable / useful findings minimal</td>
<td>Borderline satisfactory: Some skill, some reliable / useful findings likely</td>
<td>Consistent skill, manoeuvres likely to provide reliable / useful information</td>
<td>Consistent skill, manoeuvres performed will elicit reliable / useful information</td>
</tr>
</tbody>
</table>

#### Rapport with person

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condescending, offensive, judgmental</strong></td>
<td>Minimal courtesies only</td>
<td>Borderline unsatisfactory</td>
<td>Borderline satisfactory</td>
<td>Polite and interested</td>
<td>Warm, polite, empathic</td>
</tr>
</tbody>
</table>

#### Information giving

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No attempt or inappropriate attempt to give information; e.g., not truthful</strong></td>
<td>Awkward and / or incomplete attempts to give information</td>
<td>Borderline unsatisfactory, somewhat at ease, attempts to give information</td>
<td>Borderline satisfactory, somewhat at ease, attempts to give information</td>
<td>Gives information easily, somewhat attentive to patient’s understanding</td>
<td>Confident and skilful at giving information, attentive to patient’s understanding, truthful</td>
</tr>
</tbody>
</table>
Examples of Rating Scales Items

**NAC Examiner Competency Descriptors**

The following are descriptors of ACCEPTABLE performance levels per competency.

### HISTORY TAKING

**Expectations:** Acquires from the patient, family or other source a chronologic, medically logical description of pertinent events. Acquires information in sufficient breadth and depth to permit a clear definition of the patient’s problem(s).

### PHYSICAL EXAMINATION

**Expectations:** Elicits physical findings in an efficient logical sequence that documents the presence or absence of abnormalities, and supports a definition of the patient’s problem. Sensitive to the patient’s comfort and modesty; explains actions to the patient.

### ORGANIZATION

**Expectations:** Approach is coherent and succinct.

### COMMUNICATION SKILLS

**Expectations:** Uses a patient-focused approach. Shows respect, establishes trust; attends to patient’s needs of comfort, modesty, confidentiality, information. Provides appropriate, clear information and confirms patient’s understanding throughout clinical encounter. Uses repetition and summarizes to confirm and/or reinforce information, and encourages questions. Shares thinking when appropriate. Asks about patient’s support system, if appropriate. If applicable, negotiates a mutually acceptable plan of management and treatment. Demonstrates appropriate non-verbal communications (e.g., eye contact, gesture, posture, use of silence).

### LANGUAGE FLUENCY

**Expectations:** Please rate the candidate’s overall speaking skills/quality of spoken English.

Speaks clearly (appropriate volume and rate) with clear pronunciation; accent did not hinder interaction. Speaks directly to person addressed using appropriate eye contact. Provides easily understood instructions, comments and questions. Uses understandable terms for body parts and functions. Uses appropriate choice of words and expressions for the context (e.g., giving bad news). Avoids the use of jargon/slang. Uses logical flow of words, phrases and sentences and appropriate verb tenses to convey intended meaning.

### DIAGNOSIS

**Expectations:** Discriminates important from unimportant information and reaches a reasonable differential diagnosis and/or diagnosis.

### DATA INTERPRETATION

**Expectations:** Appropriately interprets investigative data in the context of the patient problem.

### INVESTIGATIONS

**Expectations:** Selects appropriate laboratory or diagnostic studies to elucidate or confirm the diagnosis; takes into consideration risks and benefits.

### THERAPEUTICS AND MANAGEMENT

**Expectations:** Discusses therapeutic management (including but not limited to pharmacotherapy, adverse effects and patient safety, disease prevention and health promotion), when appropriate. Selects appropriate treatments (including monitoring, counseling, follow-up); considers risks of therapy and instructs the patient accordingly.
### APPENDIX

#### Rating Scale Resources

**NAC Examiner Rating Guidelines**

<table>
<thead>
<tr>
<th>UNACCEPTABLE as compared to a recent graduate from a Canadian medical school</th>
<th>BORDERLINE UNACCEPTABLE as compared to a recent graduate from a Canadian medical school</th>
<th>BORDERLINE ACCEPTABLE as compared to a recent graduate from a Canadian medical school</th>
<th>ACCEPTABLE as compared to a recent graduate from a Canadian medical school</th>
<th>ABOVE the level expected of a recent graduate from a Canadian medical school</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The criteria below are meant to help support and guide examiner discrimination between the rating designations (e.g., between BORDERLINE ACCEPTABLE and ACCEPTABLE).

**UNACCEPTABLE**

- History elicited (if applicable) is incomplete and unstructured
- Physical examination (if applicable) is incomplete and unstructured
- Organization & time management skills are clearly deficient
- Communication skills & language fluency are clearly deficient
- Diagnosis/Differential diagnosis (if applicable) is inappropriate
- Data interpretation (if applicable) is incorrect
- Investigation plan (if applicable) is inappropriate
- Therapeutics and management priorities were not appropriate to the patient and the clinical case

**BORDERLINE UNACCEPTABLE**

- History elicited (if applicable) is somewhat incomplete and/or unstructured
- Physical examination (if applicable) is somewhat incomplete and/or unstructured
- Organization & time management skills are somewhat deficient
- Communication skills & language fluency are somewhat deficient
- Diagnosis/Differential diagnosis (if applicable) is limited or indiscriminate
- Data interpretation (if applicable) is poor and/or insufficient
- Investigation plan (if applicable) is poor
- Therapeutics and management priorities are poorly identified and somewhat inappropriate

**BORDERLINE ACCEPTABLE**

- History elicited (if applicable) is reasonably structured and includes some of the essential elements
- Physical examination (if applicable) is reasonably structured and includes some of the essential elements
- Organization & time management skills are just adequate
- Communication skills & language fluency are just adequate
- Diagnosis/Differential diagnosis (if applicable) is just adequate and developed in a somewhat logical manner
- Data interpretation (if applicable) is somewhat correct and/or sufficient
- Investigation plan (if applicable) is just adequate
- Therapeutics and management priorities are somewhat appropriate to the patient and the clinical case

**ACCEPTABLE**

- History elicited (if applicable) includes most of the essential elements
- Physical examination (if applicable) includes most of the essential elements
- Organization & time management skills are good
- Communication skills & language fluency are good
- Diagnosis/Differential diagnosis (if applicable) is logical and demonstrates a good understanding of the presentation
- Data interpretation (if applicable) is succinct and correct
- Investigation plan (if applicable) is appropriate and demonstrates a judicious choice of resources
- Therapeutics and management priorities were appropriate to the patient and the clinical case

**ABOVE**

- Knowledge, skills and clinical judgment clearly exceed the above criteria for an ACCEPTABLE candidate
STATION 99 – INFORMATION FOR EXAMINERS

CANDIDATE INSTRUCTIONS: Désirée Cantin, 28 years old, presents to your office today with knee pain.

IN THE NEXT 8 MINUTES:

- Obtain a focused and relevant history of the presenting problem.
- Conduct a focused and relevant physical examination.

As you proceed with the physical examination, explain to the examiner what you are doing and describe any findings.

After the 8-minute warning signal, the examiner will ask you questions related to this patient.

EXAMINER ORAL PROMPTS: None

CASE-SPECIFIC GUIDELINES

Review the case specific items below prior to rating candidates. These are guidelines which list the core performance expectations of candidates in this case. The competency headings below match the competencies on the rating scale form.

› HISTORY TAKING

Acceptable candidates should elicit the majority of the following items.

- Onset, mechanism of injury
- Location
- Aggravating/allieving factors
- Radiation
- Past history of knee problems/symptoms
- Past history of trauma
- Occupation
- Lifestyle (i.e., level of physical activity)

› PHYSICAL EXAMINATION

Acceptable candidates should perform the majority of the following items.

Inspection
- Compares left and right knees
- Observes patient standing, in bare feet
- Observes gait

Palpation
- For knee effusion (both sides)
- For tenderness along joint line

Assessment of range of motion
- Active
- Passive

Assessment of strength

Assessment of stability
- ACL: Anterior drawer test (knee flexed at 90° with foot anchored)
- AND/OR Lachman test (knee flexed at 15° with thigh stabilized)
- AND/OR pivot shift test (knee extension valgus stress imposed to knee then flexing the knee)
- PCL: Posterior drawer test
- Collateral ligaments, valgus and varus stress

Assessment of menisci
- Crouch compression test OR McMurray maneuvers

› COMMUNICATION SKILLS

- Refer to competency descriptors for communication items
### Appendix

**Patient Encounter Probe (PEP)**

Example

<table>
<thead>
<tr>
<th>Q1</th>
<th>The abdominal examination of Luc Léger revealed no organ enlargement, no masses and no tenderness. What radiologic investigation would you order first to help discriminate the cause of the jaundice?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Score</strong></td>
</tr>
<tr>
<td>A1</td>
<td>Abdominal liver ultrasound</td>
</tr>
<tr>
<td></td>
<td>Ultrasound (not specified)</td>
</tr>
<tr>
<td></td>
<td>Magnetic resonance cholangiopancreatography (MRCP)</td>
</tr>
<tr>
<td></td>
<td><strong>Maximum</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q2</th>
<th>If the investigations revealed that this patient likely had a post-hepatic obstruction, what are the two principal diagnostic considerations?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Score</strong></td>
</tr>
<tr>
<td>A2</td>
<td>Pancreatic (periampullary) cancer</td>
</tr>
<tr>
<td></td>
<td>Cancer (not specified)</td>
</tr>
<tr>
<td></td>
<td>Choledocholithiasis</td>
</tr>
<tr>
<td></td>
<td>Gallstones</td>
</tr>
<tr>
<td></td>
<td><strong>Maximum</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q3</th>
<th>What radiologic test would you consider to elucidate the level and nature of the obstruction?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Score</strong></td>
</tr>
<tr>
<td>A3</td>
<td>Magnetic resonance cholangiopancreatography (MRCP)</td>
</tr>
<tr>
<td></td>
<td>Percutaneous transhepatic cholangiography (PTC)</td>
</tr>
<tr>
<td></td>
<td>Computed tomography (CT) scan</td>
</tr>
<tr>
<td></td>
<td>HIDA scan (biliary)</td>
</tr>
<tr>
<td></td>
<td>Liver scan (technetium 99M labeled sulphur colloid)</td>
</tr>
<tr>
<td></td>
<td><strong>Maximum</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q4</th>
<th>If this patient were found to have a cancer localized to the ampulla of Vater, what single treatment would you recommend?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Score</strong></td>
</tr>
<tr>
<td>A4</td>
<td>Whipple procedure (pancreatic-duodenectomy)</td>
</tr>
<tr>
<td></td>
<td>Biliary bypass</td>
</tr>
<tr>
<td></td>
<td>Excision</td>
</tr>
<tr>
<td></td>
<td>Chemotherapy</td>
</tr>
<tr>
<td></td>
<td>Radiotherapy</td>
</tr>
<tr>
<td></td>
<td>No treatment</td>
</tr>
<tr>
<td></td>
<td><strong>Maximum</strong></td>
</tr>
</tbody>
</table>
General SP Responses in Role

If **NOT** provided in the case, these are the responses to general questions. All the responses are neutral.

- Diet: *Normal (Canada’s Food Guide)*
- Exercise: *Moderate (walk three-four times per week, x 30 minutes)*
- Medications: *No (Never)*
- Smoking: *No (Never)*
- Alcohol: *Social (e.g., wine when out for dinner once per week, 1-2 glasses)*
- Drugs: *No recreational drugs (Never)*
- Caffeine: *One coffee a day in the morning.*
- Last physical: *One year ago (normal)*
- Past medical history: *Normal (healthy)*
- Past family history: *Normal (no health issues)*
- Parents: *Healthy or died of old age (depending on SP age)*
- Siblings: *Healthy*

For women:

- Menstrual history: *Started at age 12, regular 28-day cycle, lasts four-five days,*
- Flow: *4 maxi-pads on Days 1 and 2 each. Then reduces over the last 2-3 days.*

For children:

- *All developmental milestones achieved, immunizations are up to date*

Response to any questions outside of the training information:

- “**No, I don’t have that symptom.**”
- “**No, no one in my family has that disease.**”
- “**No.**”
## APPENDIX

### OSCE Case Template

<table>
<thead>
<tr>
<th>CASE COMPONENT</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam</td>
<td>MCCQE Part II, NAC (consider purpose of exam)</td>
</tr>
<tr>
<td>Objective</td>
<td>To assess a candidate’s ability to take a history from a patient presenting with psychosis.</td>
</tr>
<tr>
<td>Assessed Discipline</td>
<td>Medicine, Surgery, Psychiatry, etc.</td>
</tr>
<tr>
<td>Assessed Skills/Roles</td>
<td>Communication, History-Taking, Physical Examination, etc.</td>
</tr>
<tr>
<td>Time Allotted for Task</td>
<td>5 minutes, 10 minutes, 11 minutes</td>
</tr>
<tr>
<td>Candidate Instructions</td>
<td>Clinical stem, task required, time allotted</td>
</tr>
<tr>
<td>Scoring Instrument</td>
<td>Checklist, rating scales</td>
</tr>
<tr>
<td>Case-Related Questions</td>
<td>Questions asked by examiner or SP; Patient Encounter Probe (PEP) questions</td>
</tr>
<tr>
<td>SP Instructions</td>
<td>Demographics, affect, physical findings, opening statements and/or questions, HPI, etc.</td>
</tr>
<tr>
<td>Room Set-up and Props</td>
<td>Hospital bed, radiographs, IV poles, reference material, etc.</td>
</tr>
<tr>
<td>Review, Revise and Pilot</td>
<td>Input from multiple content experts</td>
</tr>
</tbody>
</table>