

Medical Council  
of Canada  
Qualifying  
Examination  
(MCCQE) Part II

# 2017 MCCQE Part II Annual Technical Report



MEDICAL COUNCIL  
OF CANADA

LE CONSEIL MÉDICAL  
DU CANADA

# Table of Contents

---

<b>PREFACE</b> .....	<b>4</b>
<b>1. OVERVIEW OF THE MCCQE PART II</b> .....	<b>5</b>
<b>2. EXAM DEVELOPMENT</b> .....	<b>6</b>
2.1 Exam specifications .....	6
2.2 Case writing .....	7
2.3 Case review and piloting .....	8
2.4 Exam format .....	9
2.5 Content validity .....	9
2.6 Exam scoring .....	11
<b>3. CANDIDATE ORIENTATION, EXAM ADMINISTRATION AND TRAINING</b> .....	<b>12</b>
3.1 Candidate registration and eligibility .....	12
3.2 Candidate orientation .....	12
3.3 Exam administration .....	13
3.4 Exam administrative staff .....	13
3.5 Standardized Patient training .....	14
3.6 Deputy Registrars .....	14
3.7 Chief Examiners .....	15
3.8 Physician Examiner recruitment and training .....	15
<b>4. MCCQE PART II SCORING</b> .....	<b>17</b>
4.1 Standard quality assurance and quality control procedures .....	17
4.2 Exam result approval .....	18
4.3 Exam result reporting .....	18
4.3.1 Scale score .....	19
4.3.2 Pass/fail status .....	22
4.3.3 Domain subscores .....	22
<b>5. PSYCHOMETRIC RESULTS</b> .....	<b>23</b>
5.1 Scale scores .....	25
5.2 Estimates of reliability and classification decisions .....	27
5.2.1 Cronbach's Alpha .....	27
5.2.2 Standard error of measurement .....	28
5.2.3 Decision consistency and decision accuracy .....	28
5.3 OSCE station statistics .....	29
5.4 Examiner analyses .....	31
5.5 Domain subscore profiles .....	32
5.6 Historical pass rates .....	34
<b>6. REFERENCES</b> .....	<b>37</b>
<b>APPENDIX A: Quality control – MCCQE Part II results</b> .....	<b>38</b>
<b>APPENDIX B: Patient interaction rating scales</b> .....	<b>43</b>
<b>APPENDIX C: Example SOR and SFR</b> .....	<b>47</b>

## List of Tables and Figures

---

<b>Table 1:</b>	Exam specifications for the MCCQE Part II .....	6
<b>Table 2:</b>	Exam specifications for the MCCQE Part II (domain definitions).....	7
<b>Table 3:</b>	Sampling of OSCE exam specifications for each test form .....	10
<b>Table 4:</b>	Sampling of OSCE subscores for each MCCQE Part II test form .....	10
<b>Table 5:</b>	MCCQE Part II reporting domains.....	11
<b>Table 6:</b>	Physician examiner recruitment guidelines .....	16
<b>Table 7:</b>	Number and percentage of candidates for the MCCQE Part II by PGT group .....	24
<b>Table 8:</b>	Summary statistics of scale scores by form for each administration .....	25
<b>Table 9:</b>	Reliability estimates, standard errors of measurement, decision consistency and accuracy by form for each administration .....	29
<b>Table 10:</b>	Summary statistics for OSCE stations for each administration.....	30
<b>Table 11:</b>	2013 – 2017 pass rates by total reference group.....	34
<b>Table 12:</b>	Spring 2013 – 2017 pass rates by total reference group .....	35
<b>Table 13:</b>	Fall 2013 – 2017 Pass rates by total reference group .....	36
<b>Figure 1:</b>	Scale score distribution for spring and fall for total candidates and by PGT group .....	26
<b>Figure 2:</b>	Scale score distribution for spring and fall for CMG-CPG first-time test takers, by PGT group .....	27
<b>Figure 3:</b>	Domain subscore for spring 2017 MCCQE Part II candidates.....	33
<b>Figure 4:</b>	Domain subscore profile for fall 2017 MCCQE Part II candidates .....	33

## Preface

---

This report summarizes the exam development, exam administration, scoring and psychometric activities of the Medical Council of Canada Qualifying Examination (MCCQE) Part II and candidate performance on the exam in 2017. Sections 1 to 4 describe the exam's purpose, format, content development, administration, scoring and score reporting. These sections also provide validity evidence in support of score interpretation, reliability and errors of measurement, and other psychometric characteristics. Section 5 summarizes candidate performances for the two administrations in 2017 and includes historical data for reference purposes. The report is intended to serve as technical documentation and reference material for members of the Central Examination Committee (CEC), test committee members and Medical Council of Canada (MCC) staff.

## 1. Overview of the MCCQE Part II

---

The purpose of the Medical Council of Canada Qualifying Examination (MCCQE) Part II is to assess the competence of candidates, specifically the knowledge, skills, and attitudes essential for medical licensure in Canada, prior to entry into independent clinical practice.

The exam is a 13-station objective structured clinical examination (OSCE) that focuses on the assessment of data gathering, physical exam skills, communication skills, and considerations of cultural communication, legal, ethical and organizational (C<sup>2</sup>LEO) aspects of the practice of medicine. The exam consists of a series of 12 stations that count towards the candidate's total score and one pilot station that does not count. At each station, a brief, written statement introduces a clinical problem and directs the candidate to appropriately examine a standardized (simulated) patient (SP) and to perform activities such as obtaining a focused history, conducting a focused physical exam or assessing and addressing the patient's issues. Candidates may be asked to answer specific questions related to the patient, interpret X-rays or the results of other investigations, make a diagnosis and/or write admission orders. The MCCQE Part II includes problems in medicine, pediatrics, obstetrics and gynecology, preventive medicine and community health, psychiatry, surgery and similar disciplines considered essential for competence in general medicine and health care.

Candidates are eligible to challenge the MCCQE Part II after their medical degree has been successfully source verified, they have passed the MCCQE Part I and successfully completed a minimum of 12 months of postgraduate clinical medical training (PGT) or osteopathic postgraduate training on or before the deadline of June 30 for the spring exam of the same calendar year or December 31 for the fall exam of the same calendar year. The exam is offered twice per year, once in the spring (May) and again in the fall (October). It is scored by physician examiners (PEs) and the performance standard, also known as a pass score, reflects a candidate who is minimally competent to enter independent practice.

The Centralized Examination Committee (CEC) is responsible for the overall content and quality of the examinations and the approval of results.

## 2. Exam development

---

### 2.1 Exam specifications

Exam specifications were developed for the MCCQE Part II and approved by the OSCE Test Committee in 2004. The specifications seen below outline the domains and disciplines of the OSCE stations to ensure that similar content is measured on each of the exam forms. Creating an exam form that meets exam specifications ensures that candidates are measured on similar content and can be compared fairly from one exam to the next.

Table 1 provides the exam specifications and constraints including domain and discipline for the MCCQE Part II.

**Table 1:** Exam specifications for the MCCQE Part II

		<i>Stations</i>
<b>Domain</b>	Counselling/education	<b>2</b>
	History	<b>4 or 5</b>
	Management (ER)	<b>2</b>
	Physical exam	<b>4 or 3</b>
	<i>History + PE when possible</i>	<b>(2)*</b>
<b>Discipline</b>	Medicine	<b>3</b>
	OBGYN	<b>2</b>
	Pediatrics	<b>2</b>
	Psychiatry	<b>2</b>
	Surgery	<b>3</b>

\* *When possible, two combined history-taking and physical exam stations*

Patient interaction is not a separate domain but is integrated within most stations, except for some acute care stations.

Table 2 provides the definition for each domain outlined in Table 1 along with a description of what candidates are given credit for in each domain.

**Table 2:** Exam specifications for the MCCQE Part II (domain definitions)

Counselling/ Education	History	Management (ER)	Physical Exam	Patient Interaction
Responding to a patient problem that includes a counselling challenge	Taking an organized and focused history	Managing an acute patient problem or trauma	Conducting a focused physical exam	
Credit is given for: <ul style="list-style-type: none"> <li>◦ Eliciting key information about the patient's problem</li> <li>◦ Providing information</li> <li>◦ Understanding how the patient perceives the problem</li> <li>◦ Advising the patient</li> <li>◦ Recommending follow-up</li> </ul>	Credit is given for: <ul style="list-style-type: none"> <li>◦ Demonstrating an understanding of the most likely differential diagnosis and of the urgency of the problem</li> </ul>	Credit is given for: <ul style="list-style-type: none"> <li>◦ Setting clinical priorities</li> <li>◦ Eliciting critical information</li> <li>◦ Conducting physical diagnostic maneuvers</li> <li>◦ Ordering investigations</li> <li>◦ Starting initial treatment</li> <li>◦ Suggesting follow-up</li> </ul> Nurses are sometimes available to order tests or perform procedures requested by candidates	Credit is given for: <ul style="list-style-type: none"> <li>◦ Conducting relevant physical diagnostic maneuvers</li> <li>◦ Recognizing the level of urgency of the problem</li> <li>◦ Performing maneuvers satisfactorily</li> <li>◦ Reporting negative or positive findings</li> </ul> Standardized patients are used, not models	Patient interaction is not treated as a separate domain, but is integrated within most stations  Some acute care stations are the exception

## 2.2 Case writing

Content for the MCCQE Part II is developed by panels of physicians along with experts in medical education and assessment from across Canada. A thorough process is followed to ensure that issues related to content, feasibility, authenticity and reproducibility are addressed

early in case development. MCCQE Part II case authors develop content that reflects the MCC's expectations of competent physicians in independent practice in Canada.

Case authors create stations that deal with data gathering, diagnostic, clinical problem-solving and the principles of management that are applicable, in whole or in part, to the range of clinical situations commonly encountered by physicians.

Case authors first develop the information given to candidates prior to entering an OSCE station. This includes the required clinical task and the time allotted to candidates to perform this task. The authors then develop instruments to score the candidate's performance. Scoring instruments may consist of a checklist, rating scales, and oral or written questions. Authors also create any supporting materials, such as an X-ray or a patient chart, relating to tasks that candidates might have to complete prior to seeing the patient and/or during the patient encounter.

In addition, case authors compile information for SPs and SP trainers. To ensure that SPs and SP trainers understand the problem from the patient's perspective, case authors provide extensive information on the presenting problem, the patient's behaviour, appearance and affect, the physical findings to be portrayed as well as their appropriate medical and social history. Specific instructions about the room set-up, props, and required equipment are also provided by the authors.

## 2.3 Case review and piloting

Case development is an iterative process that requires careful thought, review, and revision. Once a case is created, its content is reviewed by clinical colleagues, test committee members, and MCC staff to improve the quality of the case and identify potential problems early in the process. For instance, role-playing the case allows authors to identify necessary information that may be missing from the SP script. Following an extensive review process, a case is piloted during a live exam to determine how well it functions both logistically and psychometrically.

The MCC is currently developing stations required to meet the domains outlined in the new Blueprint. In 2017, 14 new cases were piloted. More information about MCC's new Blueprint can be found in a report called [\*Blueprint Project: Qualifying Examinations Blueprint and Content Specifications\*](#) (Medical Council of Canada, 2014).



## 2.4 Exam format

The MCCQE Part II is a two-day examination that consists of a series of clinical stations, including pilot stations that do not count towards a candidate's total score. In 2017, candidates attempting the MCCQE Part II completed an exam form composed of eight scored, 10-minute encounter stations on Day 1 (Saturday) and four scored couplet stations on Day 2 (Sunday). Couplet stations consist of an encounter with a patient, either preceded by a task such as reading a chart or followed by a written assessment relating to the patient encounter. In 2017, the couplet stations consisted of six-minute encounters and six-minute tasks.

At each station, a brief written statement introduces a clinical problem and directs the candidate to appropriately examine an SP as well as perform activities such as obtaining a focused history, conducting a focused physical exam or assessing and addressing the patient's issues.

Standardized administration, PE/SP training, and analytic procedures are followed to ensure that the candidate's scores are comparable over time. Detailed steps of the analytic procedures are described in Appendix A: Quality Control – MCCQE Part II Results (MCC, 2017).

## 2.5 Content validity

Measuring how well an exam form matches the exam specifications is one piece of validity evidence supporting valid score interpretations and arguments for the intended purpose of the exam (Kane, 2006, 2013). This section highlights the test specifications and how well each exam form measures the exam specifications.

The MCCQE Part II Test Committee works with MCC staff to select and approve the OSCE stations for a given exam form. The exam forms are drafted by the MCCQE Part II Test Development Officer in accordance with exam specification criteria. The MCCQE Part II Test Committee then reviews the exam forms, including individual OSCE cases, to ensure that exam specification criteria have been met and that content is at the appropriate assessment level. This specifically refers to the knowledge, skills, and attitudes essential for medical licensure in Canada prior to entry into independent clinical practice. The MCCQE Part II Test Committee approves the final version of the content for each exam form. Table 3 shows the sampling of test specification characteristics for each of the MCCQE Part II forms administered in 2017. The "Recommended" column specifies the desired number of cases for each exam form per domain and discipline (as

shown in Table 1). One exam form was administered in the spring (May) and the other in the fall (October).

**Table 3: Sampling of OSCE exam specifications for each test form**

Domain	Recommended	Spring	Fall
Counselling	2	3	3
History	4-5	5	4
Combined History/PX	(2)(a)	2	3
Management	2	2	2
Physical exam	3-4(c)	1	2
Discipline	Recommended	Spring	Fall
Medicine	3	4	3
OBGYN	2	2	2
Pediatrics	2	2	2
Psychiatry	2	2	2
Surgery	3	2	3
Gender(b)		Spring	Fall
<i>As balanced as possible, at least one patient per age group</i>		M = 7	M = 6
		F = 5	F = 6
		Either M or F = 0	Either M or F = 0

(a) Combined Hx/Px station may replace a history of physical exam station

(b) Gender of SP may not necessarily be the gender of the actual patient

(c) Physical exams may be combined with management stations

Table 4 shows the frequency (as a percentage) of each of the four reporting domains sampled across the 12 scored stations for each exam form in 2017. Some domains were sampled more frequently than others. Each domain is sampled a different number of times, with some being measured across all stations and others across a smaller number of stations.

**Table 4: Sampling of OSCE subscores<sup>1</sup> for each MCCQE Part II test form**

Domain <sup>1</sup>	Target	Spring	Fall
<b>C<sup>2</sup>LEO</b>	>10%	15%	14%
<i>Data acquisition</i>		39%	38%
<i>Patient/physician interaction</i>		20%	22%
<i>Problem-solving/decision making</i>		26%	26%

<sup>1</sup> Domain subscores refer to the domains reported to candidates in their Supplemental Feedback Report (SFR).

## 2.6 Exam scoring

Each OSCE station item is assigned to one of four reporting domains: C<sup>2</sup>LEO, data acquisition, patient/physician interaction and problem-solving/decision making. PEs use checklists, oral question items, written items and rating scales<sup>2</sup> to score the candidate's proficiency on each station. Different PEs evaluate candidates in each of the stations, and one set of station scores per candidate is collected. The ratings provided by each PE are used to calculate all scores.

Table 5 lists each domain along with a description of the domains and types of items that measure each construct.

**Table 5: MCCQE Part II reporting domains**

<b>DOMAINS</b>	<b>SAMPLING</b>
<b>C<sup>2</sup>LEO</b>	The C <sup>2</sup> LEO content is included in the checklist items, interaction rating scales, oral questions and/or written questions across stations measuring that domain
<b>Data acquisition</b>	The data acquisition content is included in the items in the clinical stations related to history taking and physical examination
<b>Patient/physician interaction</b>	The patient/physician interaction content is included in the rating scale items across most stations
<b>Problem-solving/ Decision making</b>	The problem-solving and decision-making content is included in items related to the oral questions and from relevant items within the management stations and the written stations

In almost all stations, the PE also scores selected rating scale items related to the candidate's interactions with the patient. A complete list of the rating scales used in the MCCQE Part II can be found in Appendix B.

---

<sup>2</sup> Rating scales are six-level Likert items with scores ranging from zero to five. They are typically referred to as patient interaction rating scale items. Most stations include rating scales. The number of items and the specific items vary, depending on the patient problem and task for a station.

## 3. Candidate orientation, exam administration and training

---

### 3.1 Candidate registration and eligibility

To be eligible to take the MCCQE Part II, candidates must pass the MCCQE Part I, their final medical degree diploma must be successfully source verified, and they must have successfully completed a minimum of 12 months of postgraduate clinical medical training (PGT) or osteopathic postgraduate training (PGY-2+); or completed 12 months of postgraduate clinical medical training (PGT) on, or before the deadline of June 30 for the spring exam of the same calendar year or December 31 for the fall exam of the same calendar year (PGY-1). The MCC implemented capacity limits to registration for the spring and fall administrations beginning in fall 2015 due to the demand in particular sessions exceeding the number of tests spots and the concern that increased demand would affect the quality of the examination administration. Additionally, implementing capacity limit was a way to ensure an adequate sample of candidates to conduct linking of MCCQE Part II total scores. At each administration, PGY-2+ candidates are given priority on available examination spots. If additional spaces are available, PGY-1 candidates are provided the opportunity to apply for the remaining spots.

For the spring and fall 2017 exams, candidates with a minimum of 12 months of training were invited to add their names to a MCCQE Part II pre-application list through their [physiciansapply.ca](http://physiciansapply.ca) account during a specific pre-application period. When the pre-application period ended, candidates were randomly selected from the pre-application list based on available exam capacity. For the fall 2017 exam, there was not enough exam capacity to accommodate all PGY-2+ candidates and therefore no PGY-1 candidates were selected.

### 3.2 Candidate orientation

The MCC provides candidates with detailed information about the MCCQE Part II on its website. Topics include what to expect on examination day, scoring, results and application information as well as an online orientation presentation. Candidates must also participate in a mandatory orientation given by senior site staff on each exam day before the exam begins. These sessions provide candidates with:

- Information on the personal belongings that a candidate can and cannot bring to the exam

- Information on how a candidate may use their booklet and their bar code labels
- An overview of the length and number of stations and how they will rotate through the stations
- Information on the timing of the stations and how the signal system functions
- Guidance on how to interpret the candidate instructions
- Information on how they will be assessed by the PEs and when a PE may intervene
- Information on available medical equipment
- A description of how the candidate should interact with SPs
- An overview of the short-answer write-ins and the multiple-choice written stations
- Instructions on exam security and how to ask for assistance
- Information about the MCC's conflict of interest policy, confidentiality, sequestration, expected candidate behaviour, and how to report concerns on exam day and to the MCC after the exam

### 3.3 Exam administration

In 2017, 4,640 candidates participated in the MCCQE Part II across two administrations and two exam forms. Exams were administered in English in Calgary, Edmonton, Halifax, Hamilton, Kingston, London, Montreal, Ottawa, Saskatoon, St. John's, Sudbury, Toronto, Vancouver, Victoria and Winnipeg. The exam was also offered in French in Montreal, Quebec City and Sherbrooke.

### 3.4 Exam administrative staff

Each partner site is responsible for hiring and supervising administrative staff. They work with the MCC to safeguard exam materials and to ensure that all people involved in the exam (site administrators, SP trainers, SPs, Chief Examiners (CE), PEs, exam day staff, caterers, etc.) perform to standard.

MCC personnel oversees site staff on exam days across the country in person, by telephone, and via electronic communication and works a hotline on exam days.

Every two years, the MCC hosts a face-to-face meeting for all MCCQE Part II site administrators to review administrative aspects of the examination, troubleshoot site-specific challenges such as PE recruitment and enhance collaboration between sites.

### 3.5 Standardized Patient training

Each site is responsible for hiring and supervising the SP trainers who oversee the SPs and assure the quality of their standardized performances on exam day(s). SPs are trained at each site using MCCQE Part II training materials provided by the MCC. Training support is provided centrally by MCC staff, primarily by the MCCQE Part II training officer.

Every two years, the MCC hosts a face-to-face meeting for all MCCQE Part II SP trainers to review SP training aspects of the exam, troubleshoot site-specific challenges such as SP recruitment and to enhance collaboration between sites.

### 3.6 Deputy Registrars

The Deputy Registrar (DR), on behalf of the Dean, is responsible for the administrative and financial operations of the exam centre and the local administration of the MCCQE Part II. The DR is also responsible for:

- Ensuring that an office is established for the team, adequate office space and furnishings are available, including computer equipment, phones and copiers, and that security precautions are taken
- Selecting and supervising centre personnel, especially the exam administrator as well as hiring senior site staff, a task usually performed in consultation with the manager of the MCCQE Part II
- Providing leadership and support to senior site staff for planning and teamwork
- With assistance of the CE, recruiting physicians to serve as examiners, assigning them to their stations and conducting pre-exam orientation sessions for both first-time and experienced examiners
- Delivering orientation to CEs
- Assisting with dry runs for the SPs for quality assurance, a role usually shared with the CE
- Ensuring the administration runs smoothly, especially in the following areas:

- Checking props and room setup
- Supervising and supporting staff throughout the day
- Ensuring the examination centre is securely closed at the end of the exam day
- Reviewing all incident reports and ensuring appropriate action is taken and documented
  - Incident reports are completed to report the details of unusual events or breaches in procedures to the MCC; these reports may be completed by candidate(s), PEs, SPs, SP trainers, the DR, the CE, the Site Administrator or site staff
- Liaising with MCC before, during and after the exam
- Overseeing the preparation of a financial report, and interim financial report, if necessary, and the DR's report for the exam session
- If requested, reviewing and commenting on training materials, cases and other documentation before deadlines

### 3.7 Chief Examiners

The CE's role depends on the size of the centre and how the DR assigns responsibilities. Generally, the CE should assist with SP dry runs, participate in some of the SP training sessions and assist in examiner recruitment and staff training as needed. All SPs take part in a dry run prior to exam day with the CE playing the candidate's role to ensure they are ready to perform their roles. An SP who is deemed as not being ready is re-trained or replaced.

### 3.8 Physician Examiner recruitment and training

Table 6 presents the guidelines used to recruit PEs for operational or pilot stations.

---

<sup>3</sup> An incident report is a form that is completed by candidates or any member of the exam site staff to report an incident to the MCC. An incident is any irregular event that occurs before, during or after the exam. Examples include security breaches, late-arriving candidates, signal errors, SP portrayal errors, etc. Incident reports are to be returned to the MCC with the score sheets. The MCC reviews and evaluates all incident reports. Incidents affecting candidate's exam status are presented to the CEC.

**Table 6: Physician examiner recruitment guidelines**

GUIDELINES FOR PHYSICIAN EXAMINER RECRUITMENT	
<p>Must have the Licentiate of the Medical Council of Canada (LMCC) and must provide their LMCC registration number.</p>	
<p>Physicians in Quebec who do not have their LMCC will be accepted as examiners under the following conditions:</p> <ul style="list-style-type: none"> <li>• Non-Licentiate examiners must not exceed 50 per cent of the cohort of examiners assigned to a given session</li> <li>• Non-Licentiate examiners must be senior faculty members</li> <li>• Non-Licentiate examiners must be certified by the College of Family Physicians of Canada (CFPC), RCPSC or the Collège des médecins du Québec (CMQ)</li> </ul>	
<p>Should have the ability and stamina for the task (for example, hearing loss would hamper an examiner's ability to accurately score)</p>	
<p>PEs participating in any MCCQE Part II preparatory course(s) are not eligible to be examiners for the MCC for a minimum of three years</p>	
<p><b>Operational stations:</b> <i>PEs meeting these criteria may examine in both operational and pilot stations</i></p>	<ul style="list-style-type: none"> <li>• Must be two years post-LMCC</li> <li>• Must have at least two years in independent practice</li> <li>• Cannot be residents</li> <li>• Examiners must hold an unrestricted licence and currently be practising medicine</li> <li>• May be community physicians</li> <li>• Can be a fellow but must meet all other criteria, except criteria #2</li> </ul>
<p><b>Pilot stations:</b> <i>PEs meeting these criteria may examine in pilot stations only</i></p>	<ul style="list-style-type: none"> <li>• Resident physicians must be PGY-4 level or higher OR have CCFP certification</li> <li>• Must have recent experience supervising clerks and/or PGY-1s, and/or experience as an examiner at this level of training (for example, a resident who worked as an examiner for a University OSCE, or other similar OSCEs, but not the MCC OSCEs)</li> </ul>

The MCC provides an exam day PE orientation given by the sites' CEs, DR and/or senior site staff. Before every exam, all PEs must also participate in an online PE orientation that provides information on enhancing the standardization of PE scoring, including a practice scoring session and a guided discussion.



## 4. MCCQE Part II scoring

---

This section describes quality assurance and quality control procedures relating to the scoring of candidates on the MCCQE Part II, what scores are reported and how they are calculated.

### 4.1 Standard quality assurance and quality control procedures

To ensure the accuracy and integrity of the candidate's exam day electronic records, quality assurance (QA) steps are performed as outlined below.

PEs complete score sheets for every candidate seen in their OSCE stations. These forms are scanned at each exam site and transmitted securely to the MCC where MCC staff import the forms into Cardiff TeleForm® software where they are reviewed. Scanning anomalies are identified (for instance, a non-scannable candidate barcode label, a PE's pencil marks that are too faint, missing sheets due to candidates that do not complete the exam on both days) and the requisite corrections are made to the electronic data records. Data is imported electronically into a scoring application to generate a list of all candidates whose total score is close to (~3 per cent below and ~3 per cent above) the pass score. We call this group the selected candidate group. Paper copies of the score sheets for this selected candidate group are visually reviewed. Some examples of checks include:

- Number of selections per question for extended match questions where candidates can bubble a number of correct answers
- Notes by PEs for oral or SP questions
- Confirmation of missing data for oral questions or rating scales
- Flags for lapses in patient safety or professional behaviours
- Confirmation of scores and bubbles for written items
- Verification of raw score points

Any differences are corrected in the electronic data files to reflect the paper score sheets. There are typically 100 to 120 corrections made on each exam. The updated electronic files are then re-imported into the scoring application that is used to create the scale scores for all candidates. All scores are calculated in parallel using the *Statistical Analysis System (SAS®)* and compared to the results from the scoring application. All values must match before results are released to candidates.

## 4.2 Exam result approval

The results for each administration of the MCCQE Part II are reviewed by the CEC. The CEC approves the release of results after each administration, including special cases. Once the CEC has approved the results, they are imported and released to candidates.

Exam results that meet the following criteria are automatically approved by the CEC: (1) station and item level p-values between 0.10 and 0.90, (2) Station Total Correlation (STC) greater than 0.30 and Item Total Correlation (ITC) greater than 0.05, and (3) decision consistency and accuracy values greater than 0.80. If the p-values, STC or ITC are not within the approved values the content is reviewed by subject matter experts to ensure that the content is still appropriate. Automatically approved exam results are reported to candidates once all of the quality assurance processes are completed. Special cases are not automatically approved.

When an incident occurs that may impact a candidate's performance on exam day, it is presented to the CEC as a special case. The CEC determines the severity of the impact and decides if any changes should be made to the candidate's exam results. Depending on the nature of the incident the CEC may decide to remove one or more items from a candidate's exam, award the candidate a "No Standing", or a "Denied Standing".

A "No Standing" indicates that procedural irregularities in the examination process may have materially affected the performance of the candidate and/or may have prevented a reliable assessment of the candidate's knowledge and abilities.

A "Denied Standing" indicates that a candidate has been found to have committed an infraction related to the MCC's examination process and/or breached the confidentiality of the examination. Candidates that are awarded a "Denied Standing" may be denied entry to one or more future examinations of the MCC.

## 4.3 Exam result reporting

Approximately seven weeks after the last day of the exam session, the MCC issues a Statement of Results (SOR) and Supplemental Feedback Report (SFR) to each candidate through their physiciansapply.ca account (samples from fall 2016 administration are shown in Appendix C). The SOR includes the candidate's final result and total score, as well as the pass score.

Additional information about the total and subscores and comparative information is provided on the SFR. The total score is reported on a standard-score scale ranging from 50 to 950. In contrast, the score profile in Figure 1 of the SFR example displays a candidate's domain subscores that indicate a candidate's relative strengths and weaknesses in four areas. It is important to note that, because subscores have fewer items, there is less measurement precision. Subscores are provided to individual candidates for feedback only and are not meant to be used by organizations for selection. The following sections outline the steps in creating the results reported to candidates.

### **4.3.1 Scale score**

---

Deriving the scale score for the MCCQE Part II involves three steps.

#### **Step 1: Calculate total scores**

The first step in obtaining a scale score is to calculate the total score for each candidate. To do so, a station score is calculated for each station using the following formula:

$$\text{station score} = \frac{\sum_i \text{score}_i}{\sum_i \text{max\_score}_i} * 100$$

where the numerator is the sum of each candidate's scores on each item  $i$  for that station and the denominator is the sum of the maximum score for each item for that station. For example, a station with several checklist items, an oral question, and several rating scales could result in the following score:

$$\text{station score} = \frac{\text{sum of items}=60}{\text{sum of max score of items}=67} * 100=89.55$$

The station scores are then used to calculate the total score for each candidate using the following formula:

$$\text{total score} = (\text{sum of 12 station scores})/12$$

Since station scores are based on the sum of the candidate's items for that station, missing data needs to be taken into account so that it does not negatively impact a candidate's score. Missing data occurs when the PE does not provide a rating for some of the oral questions or rating scales for a given candidate on the scannable score sheet. When oral questions or rating scales are not

scored by the PE, the station score is based on the items that are provided by the PE. Extending the example above, a station that has one missing rating scale item with a maximum score of five would result in the following score:

$$\text{station score} = \frac{\text{sum of candidate items}=60}{\text{sum of max score of items}=62} * 100 = 96.77$$

The station score would have been 89 per cent if the missing rating scale was treated as zero and the adjustment not applied. However, to be fair to the candidate, we exclude the missing rating scale from the calculation of the station score.

### ***Step 2: Link MCCQE Part II scores to base test form – spring 2015***

For each MCCQE Part II administration, one test form is used to measure the exam specifications (as described in the Exam specifications section). The MCC staff and the MCCQE Part II Test Committee work in collaboration to select the 10-minute OSCE stations and the couplet OSCE stations that best represent the Blueprint. One possible adverse effect of having different test forms with different stations across administrations is that one test form may be more difficult than the other. However, the process of linking total scores statistically takes into account differences in test form difficulty and adjusts total scores so that all scores are on the same metric. Linking is also a way of applying the same pass score to candidates who take different test forms.

The next step in obtaining scale scores for all candidates is to link scores through the common stations to the base form. For the 2017 spring and fall MCCQE Part II, linking occurred to place the candidate's scores on the same metric as the base form from the spring 2015 MCCQE Part II administration.

One method of linking test forms is to have a subset of the content appear identically across test forms; this is called a non-equivalent anchor test or NEAT design. The subset of content that is presented identically is called an anchor set. The rule of thumb for determining the number of items in an anchor set for a multiple-choice exam is 20 per cent of the total test or 20 items, whichever is greater, to ensure that the anchor set is representative of the total test in terms of content and difficulty. Since the MCCQE Part II is an OSCE with a small number of stations (less than 20), we use a 30 per cent rule. The anchor set is used to statistically estimate the overall ability of candidates, taking each test form and the difficulty of the test forms into account.

For the MCCQE Part II 2017 spring and fall test forms, the anchor set was based on four stations. A reference group of Canadian medical graduate and Canadian postgraduate first-time test takers was used for all linking calculations. In all linking steps, the Levine's observed score method was employed. Full details of the method can be found in *Testing equating, scaling, and linking: Methods and Practice* (2nd Edition) by Kolen and Brennan (2004).

### **Step 3: Transform linked scores to scale scores**

Once the first two steps are complete, the linked scores are transformed to scale scores ranging from 50 to 950 for reporting purposes. Using the spring 2015 MCCQE Part II results from all candidates, the new scale was established to have a mean of 500 and a standard deviation of 100. This final transformation ensures that any differences in candidate scale scores, or mean scale scores and standard deviations on subsequent test forms can be directly compared to the spring 2015 MCCQE Part II results. For example, a candidate score or mean cohort increase to 600 or decrease to 450 would indicate the general performance of the candidate or group of candidates who took the subsequent test forms was higher and lower respectively.

The final transformation formula for all test forms is as follows:

$$\text{ScaleScore}_x = (\text{slope})(\text{LinkedScore}_x) + (\text{intercept})$$

Where  $\text{ScaleScore}_x$  is defined as the linear function to calculate the scale score for candidate X, where the slope is equal to 15.08 and applied to all test forms (based on the initial transformation of the spring 2015 MCCQE Part II), where the intercept is equal to -459.86 and also applied to all test forms (based on the initial transformation of the spring MCCQE Part II) and where the  $\text{LinkedScore}_x$  is the linked score for candidate X.

All scale scores are rounded to a whole number between 50 and 950. The reported scale scores as seen by candidates is this rounded value. For example, a passing candidate with a linked score of 81.25 would have a scale score of 765:

$$\text{ScaleScore}_x = (15.08) * (81.25) + (-459.86) = 765.39 \text{ rounded to } 765$$

A failing candidate with a linked score of 42.51 would result in a scale score of 181:

$$\text{ScaleScore}_x = (15.08) * (42.51) + (-459.86) = 181.19 \text{ rounded to } 181$$

### **4.3.2 Pass/fail status**

---

The pass score for this exam was set by a panel of 20 physicians who reviewed stations, content and score information and provided judgments for establishing the recommended pass score. The fall 2014 MCCQE Part II test form was used to establish the pass score and was first applied to the group of candidates in the spring 2015 administration. The pass score from the fall 2014 test form was linked to the score scale from spring 2015, using a Levine observed score linking method. The borderline group method was used for the standard-setting exercise and to calculate the pass score. Full details of the standard-setting exercise can be found in the [\*Technical Report on the Standard Setting Exercise for the Medical Council of Canada Qualifying Examination Part II\*](#) (Medical Council of Canada, 2015). The pass score of 509 was recommended by the panel of physicians for spring 2015 and approved by the CEC in June 2015. This pass score was used to assign a pass/fail status to each candidate in 2017. A score of 509 or greater is deemed a pass.

### **4.3.3 Domain subscores**

---

Domain subscore calculations are used to create the figure in the candidates' SFRs. For each domain subscore, the associated items are converted to a percentage ranging from 0 to 100, where the total number of score points obtained by a candidate is divided by the maximum score points per domain, multiplied by 100.

For example, if a candidate received scores of five, seven, eight and one on a domain with associated maximum scores of 10, 10, nine and one, the total number of score points obtained by the candidate is 21; the maximum number of score points for this domain is 30. The domain score is  $21/30 \times 100$  or 70.0. For the MCCQE Part II, there are four domain subscores that are presented to candidates in their SFRs: C<sup>2</sup>LEO, data acquisition, patient/physician interaction and problem-solving and decision making (PSDM).

Domain subscores are not used to calculate the total score or scale scores, as outlined above; therefore, domain subscores cannot be directly compared to the candidates' scale scores. Domain subscores provide some general feedback to candidates on their relative strengths and weaknesses on their performance on the MCCQE Part II.

## 5. Psychometric results

---

The data used for the aggregate analyses are the results approved by the CEC in June 14, 2017 for the spring administration and December 6, 2017 for the fall administration. In total, 4,640 candidates participated in the MCCQE Part II administered in May and October 2017.

The MCC implemented capacity limits to registration for the spring and fall 2017 administrations as mentioned previously. In addition, for the spring 2017 administration, candidate eligibility allowed residents who had approximately 10 months of PGT, called PGY-1, by the exam date to take the MCCQE Part II. Prior to 2017, candidates were only eligible if they had completed at least 12 months of PGT at the time of the exam, called PGY-2 or above, PGY-2+. In 2017, PGY-1 candidates were able to register for an exam if there were adequate spaces available for that administration (this occurred only for the spring administration in 2017).

The number of candidates for each administration is depicted in Table 7. It shows the number of candidates for the spring and fall administrations by candidate groups (for example, Canadian medical graduates (CMG), first-time test takers (1st), etc.), gender and examination language. The main reference group for this examination includes CMGs and the Canadian, postgraduate, first-time test-taker candidate group. Candidate groups shown in Table 7 are as follows:

- CMG, Canadian postgraduate, first-time test takers (CMG-CPG 1st)
- CMG, Canadian postgraduate, repeat test takers (CMG-CPG repeaters)
- CMG, international postgraduate, first-time test takers (CMG-IPG 1st)
- CMG, international postgraduate, repeat test takers (CMG-IPG repeaters)
- International medical graduate (IMG), international postgraduate, first-time test takers (IMG-IPG 1st)
- IMG, international postgraduate, repeat test takers (IMG-IPG repeaters)
- IMG, Canadian postgraduate, first-time test takers (IMG-CPG 1st)
- IMG, Canadian postgraduate, repeat test takers (IMG-CPG repeaters)

**Table 7: Number and percentage of candidates for the MCCQE Part II by PGT group**

PGT Group	Subgroups	Spring (a)		Fall (b)		Total	
	Candidate group	N	%	N	%	N	%
PGY-1	CMG-CPG 1st	280	81.4	0	0.0	280	81.4
	CMG-CPG repeaters	0	0.0	0	0.0	0	0.0
	CMG-IPG 1st	0	0.0	0	0.0	0	0.0
	CMG-IPG repeaters	0	0.0	0	0.0	0	0.0
	IMG-IPG 1st	4	1.2	0	0.0	4	1.2
	IMG-IPG repeaters	0	0.0	0	0.0	0	0.0
	IMG-CPG 1st	57	16.6	0	0.0	57	16.6
	IMG-CPG repeaters	3	0.9	0	0.0	3	0.9
	<b>Gender</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>
	Female	166	48.3	0	0.0	166	48.3
	Male	178	51.7	0	0.0	178	51.7
	<b>Language</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>
	English	270	78.5	0	0.0	270	78.5
	French	74	21.5	0	0.0	74	21.5
	<b>Total</b>	<b>344</b>	<b>100.0</b>	<b>-</b>	<b>-</b>	<b>344</b>	<b>100.0</b>
PGY-2+	<b>Candidate group</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>
	CMG-CPG 1st	884	42.5	1709	66.8	2593	55.9
	CMG-CPG repeaters	203	9.8	62	2.4	265	5.7
	CMG-IPG 1st	2	0.1	1	0.0	3	0.1
	CMG-IPG repeaters	0	0.0	0	0.0	0	0.0
	IMG-IPG 1st	438	21.1	339	13.2	777	16.7
	IMG-IPG repeaters	283	13.6	219	8.6	502	10.8
	IMG-CPG 1st	151	7.3	178	7.0	329	7.1
	IMG-CPG repeaters	119	5.7	52	2.0	171	3.7
	<b>Gender</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>
	Female	908	43.7	1320	51.6	2228	48.0
	Male	1172	56.3	1240	48.4	2412	52.0
	<b>Language</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>	<b>N</b>	<b>%</b>
	English	1731	83.2	2178	85.1	3909	84.2
	French	349	16.8	382	14.9	731	15.8
<b>Total</b>	<b>2080</b>	<b>100.0</b>	<b>2560</b>	<b>100.0</b>	<b>4640</b>	<b>100.0</b>	

(a) One "No Standing" candidate is not included in the remaining analyses

(b) Two "No Standing" candidates are not included in the remaining analyses. One candidate with more than one station missing is not included in the remaining analyses, with the exception of tables 8, and 11 through 13 where pass rates are displayed.



In the sections below, summary statistics for scale scores and pass rates are provided, as well as estimates of reliability for both scores and classification decisions. Finally, a summary of station quality and domain subscore profiles are also outlined.

## 5.1 Scale scores

Scale score summary statistics from the spring and fall MCCQE Part II that were reported to the CEC are presented in Table 8. They are based on the scale scores that include the minimum, maximum, mean and median scale scores, standard deviation and pass rates for the total group and for CMG, Canadian postgraduate (CPG), first-time test takers (main reference group).

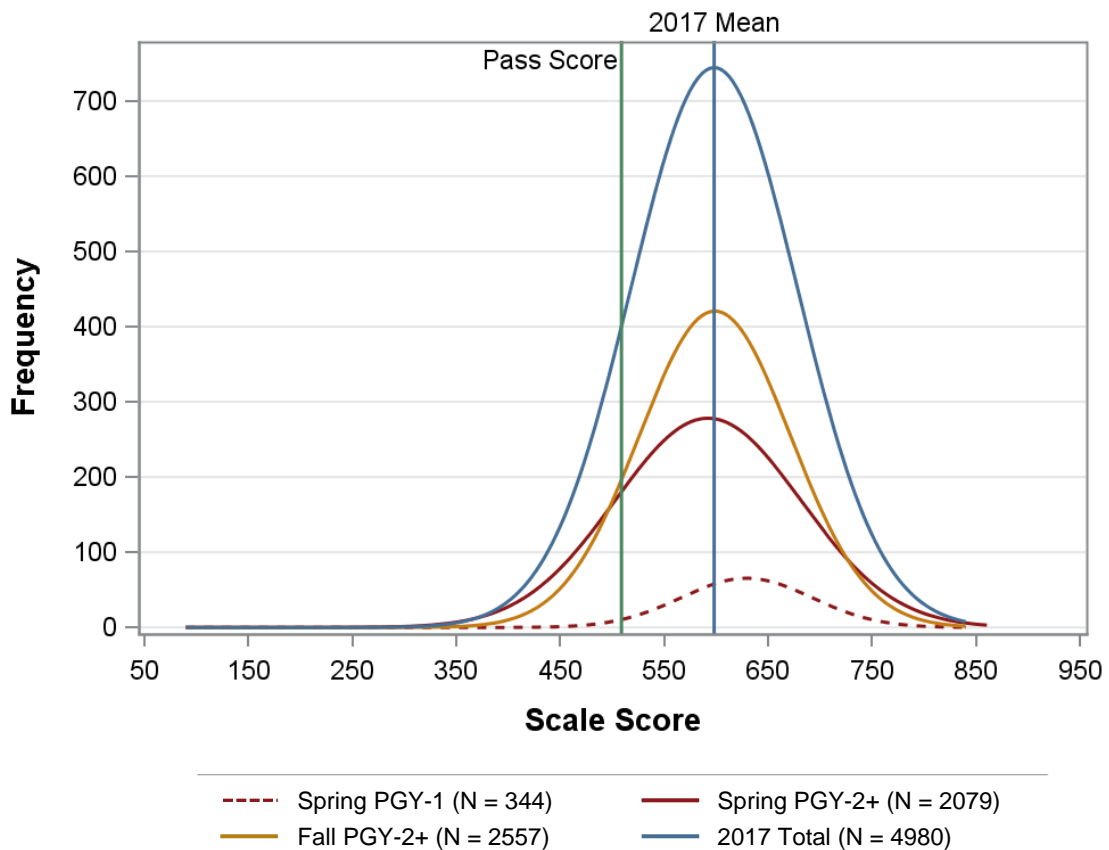
**Table 8: Summary statistics of scale scores by form for each administration**

<i>Exam session</i>	<i>Candidate group</i>	<i>PGT group</i>	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>	<i>Pass rate</i>
<b>Total</b>	Total	PGY-1	344	449	788	629.3	633.0	62.9	96.5
		PGY-2+	4636	98	817	595.9	602.0	80.7	86.4
	CMG-CPG 1st	PGY-1	280	470	788	636.7	635.0	59.2	97.9
		PGY-2+	2591	394	817	631.0	632.0	63.0	97.1
<b>Spring</b>	Total	PGY-1	344	449	788	629.3	633.0	62.9	96.5
		PGY-2+	2079	98	808	592.2	597.0	89.4	82.6
	CMG-CPG 1st	PGY-1	280	470	788	636.7	635.0	59.2	97.9
		PGY-2+	883	395	808	649.3	651.0	65.2	98.0
<b>Fall</b>	Total	PGY-1	0	-	-	-	-	-	-
		PGY-2+	2557	285	817	598.8	604.0	72.7	89.5
	CMG-CPG 1st	PGY-1	0	-	-	-	-	-	-
		PGY-2+	1708	394	817	621.5	623.0	59.7	96.7

The mean, median scale scores and pass rates vary by exam administration. The minimum, maximum and standard deviation are indicators of the variation in scale scores.

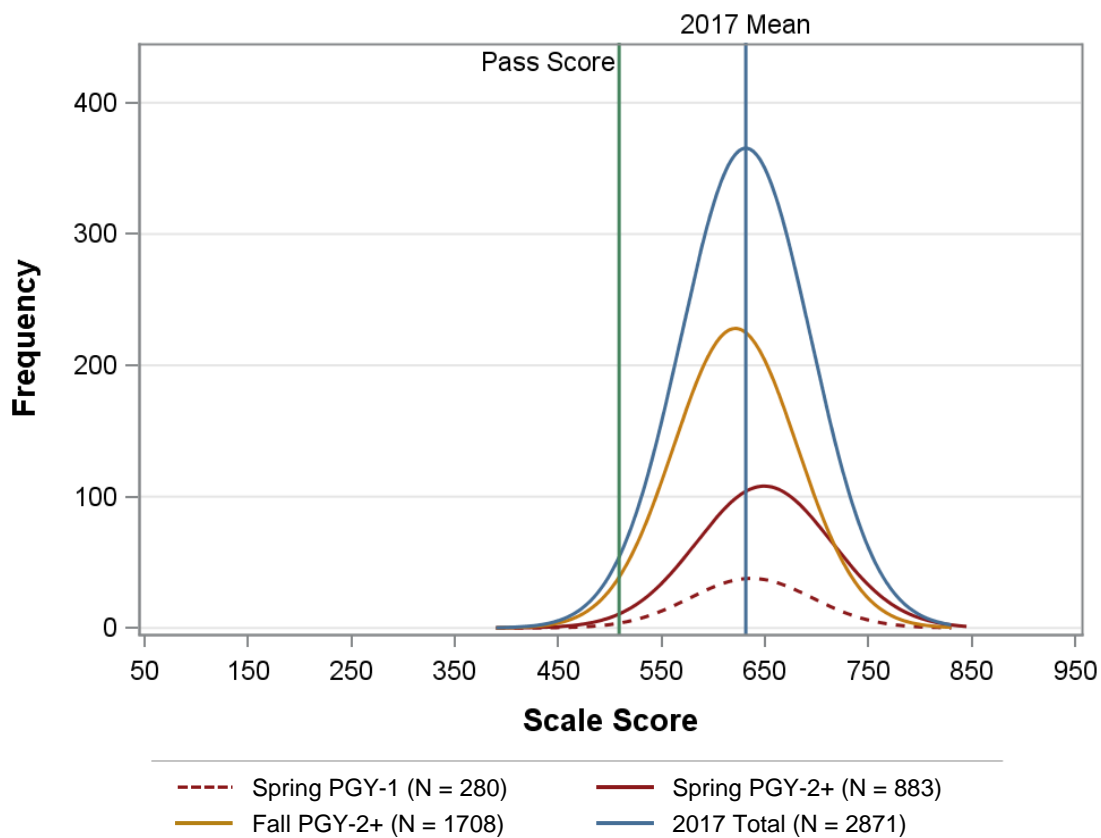
Figure 1 displays the distribution of scale scores for the MCCQE Part II candidates for spring and fall 2017, by PGT Group (such as PGY-1 and PGY-2+). Figure 1 shows that a smaller proportion of candidates were in the PGY-1 group than the PGY-2+ group and the candidates in PGY-1

group were slightly higher performing than either spring or fall PGY-2+ groups. Scale scores were generally similar for candidates who took the spring 2017 MCCQE Part II versus those who took the fall 2017 MCCQE Part II.



**Figure 1:** Scale score distribution for spring and fall for total candidates and by PGT group

Figure 2 displays the distribution of scale scores for spring and fall 2017 by PGT group for CMG-CPG 1st candidates. Figure 2 shows that a smaller proportion of candidates were in the PGY-1 group than the PGY-2+ group and the candidates in PGY-1 group were slightly lower performing than spring and higher performing than fall PGY-2+ groups. Although there were a higher proportion of CMG-CPG 1st candidates in fall 2017 the performance of candidates in the CMG-CPG 1st group across spring and fall was very similar.



**Figure 2:** Scale score distribution for spring and fall for CMG-CPG first-time test takers, by PGT group

## 5.2 Estimates of reliability and classification decisions

### 5.2.1 Cronbach's Alpha

Cronbach's alpha was used to estimate score reliability for the MCCQE Part II test forms. A score reliability estimate indicates the desired consistency (or reproducibility) of examination scores across replications of measurement (Crocker & Algina, 1986; Haertel, 2006). Scores that are highly reliable are accurate, reproducible and consistent from one testing occasion to another. In other words, if the testing process was repeated with a group of test takers, essentially the same results would be obtained. The reliability estimate is further described in Educational

Measurement by Haertel in chapter 2, section 2.4.4 (Haertel, 2006). The formula for Cronbach's alpha is:

$$\alpha \rho_{XX'} = \frac{n}{n-1} \left( 1 - \frac{\sum \sigma_{X_i}^2}{\sigma_X^2} \right)$$

where  $n$  is the number of stations,  $\sigma_{X_i}^2$  is the variance of station  $i$  scores and  $\sigma_X^2$  is the variance of total scores (Haertel, 2006, p. 74). As a general rule, reliability is desired to be greater than 0.80 on an OSCE. The reliability estimate, in conjunction with the total exam standard error of measurement (SEM), can provide further evidence of the reliability of the candidate's scale score.

### **5.2.2 Standard error of measurement**

---

The SEM provides a value within a certain confidence range (for example, 68 per cent or 95 per cent) that a candidate's observed score is expected to range if the candidate was retested over repeated exams that are similar in measuring the same test specifications. SEM values should be as small as possible so that measurement of the candidate's ability contains as little error as possible. The SEM is calculated as follows:

$$SEM = SD \sqrt{1 - \alpha \rho_{XX'}}$$

where  $SD$  is defined as the standard deviation for the total score (square root of the variance) and  $\alpha \rho_{XX'}$  is defined as the reliability estimate as shown above.

### **5.2.3 Decision consistency and decision accuracy**

---

Estimates indicating the consistency and accuracy of pass/fail decisions are important in providing validity and reliability evidence for candidate scores on one test form with possible equivalent test forms. To this end, the MCCQE Part II uses the Livingstone and Lewis (1995) procedure, where decision consistency is an estimate of the agreement between classifications on potential parallel test forms and decision accuracy is the estimate of agreement between the observed classifications of candidates and those based on their true score (observed score  $\pm$  measurement error). Ideally, both values should be high, such as 0.80 and above, supporting a reliable and valid pass/fail standing. A value of 0.80 indicates that accuracy or consistency of the decision is being met for at least 80 per cent of the candidates.

Table 9 shows the reliability estimates, the SEM and the decision consistency and decision accuracy estimates along with the associated false positives and false negative rates by test form. The estimated false positive rate indicates the expected proportion of candidates who pass based on their observed scores but who should fail based on their true ability. The estimated false negative rate indicates the expected proportion of candidates who fail based on their observed scores but who should pass based on their true ability.

**Table 9:** Reliability estimates, standard errors of measurement, decision consistency and accuracy by form for each administration

	Spring	Fall
<b>Reliability estimate</b>	0.77	0.68
SEM (score scale)	41.93	41.43
<b>Decision consistency</b>	0.89	0.90
False positive	0.06	0.05
False negative	0.06	0.05
<b>Decision accuracy</b>	0.92	0.93
False positive	0.03	0.02
False negative	0.05	0.05

It should be noted that reliability is impacted both by the amount of variability in scores amongst candidates taking a particular test form and the number of items or stations included in any given exam. It is more difficult to obtain reliability estimates above 0.80, given the restricted number of stations that can be administered in any OSCE form. The decision consistency and accuracy values should be above 0.8 for OSCEs and for both spring and fall these values were well above 0.8.

### 5.3 OSCE station statistics

Summary statistics for each of the OSCE stations by administration are provided in Table 10. The percentage of missing data, proportion correct or  $p$ -values, standard deviation and STCs are presented.

$P$ -values are the average proportion correct scores that candidates achieved on each of the stations. In general,  $p$ -values indicate station difficulty and range between 0 and 1. Station  $p$ -values that are low (<0.20) indicate a difficult station; those that are high (>0.90) indicate an easy

station. *P*-values are population dependent. That is, comparisons of *p*-values across different samples of candidates do not take into account potential differences in overall candidate ability. As such, *p*-values should not be overinterpreted or used as the only indicator of difficulty. Rather, *p*-values provide a general sense of the range of difficulty of stations on a particular test form.

Standard deviations indicate the general variability of scores on any given station. STCs are indicators of discrimination between low- and high-ability candidates for a given station. A low positive or negative STC (-0.20 to <0.30) indicates that there is a weaker relationship between the station score and the overall exam score. Along with the *p*-values, this information may be useful in flagging stations that should be reviewed by content experts and possibly removed from scoring. A moderate to high STC (>0.30) indicates that high-ability candidates are performing well on a given OSCE station. Stations with STCs that are below 0.30, as well as negative values, are flagged for review. On occasion, flagged and reviewed stations may still be included on an exam when the content is deemed relevant, important and has been verified to be correct.

**Table 10: Summary statistics for OSCE stations for each administration**

Station	Spring				Fall			
	% missing	<i>p</i> -value	SD	STC	% missing	<i>p</i> -value	SD	STC
1	0.01	0.67	0.11	0.42	0.00	0.71	0.14	0.14
2	0.02	0.68	0.12	0.45	0.01	0.68	0.11	0.35
3	0.00	0.76	0.12	0.31	0.01	0.67	0.10	0.39
4	0.04	0.71	0.14	0.36	0.01	0.75	0.10	0.39
5	0.01	0.64	0.15	0.43	0.03	0.64	0.13	0.34
6	0.02	0.75	0.11	0.41	0.02	0.67	0.10	0.30
7	0.00	0.67	0.11	0.47	0.01	0.67	0.12	0.41
8	0.02	0.61	0.12	0.38	0.01	0.66	0.13	0.40
9	0.03	0.57	0.15	0.32	0.01	0.58	0.14	0.21
10	0.01	0.65	0.11	0.43	0.04	0.68	0.14	0.29
11	0.01	0.69	0.11	0.46	0.03	0.65	0.13	0.34
12	0.01	0.71	0.11	0.43	0.02	0.67	0.14	0.30
Min	0.00	0.57	0.11	0.31	0.00	0.58	0.10	0.14
Max	0.04	0.76	0.15	0.47	0.04	0.75	0.14	0.41
Mean	0.01	0.68	0.12	0.41	0.02	0.67	0.12	0.32
SD	0.01	0.05	0.01	0.05	0.01	0.04	0.02	0.08

Table 10 shows the mean  $p$ -values for each administration. There were no stations flagged as being too difficult ( $p$ -value  $<0.30$ ) or too easy ( $p$ -value  $>0.90$ ). The standard deviation for each test form also indicates there was reasonable variation in station scores. For the STCs, most values were within acceptable ranges (STC  $>0.30$ ). These results suggest that these OSCE stations discriminated well between low- and high-ability candidates. Stations with an STC  $<0.30$  were reviewed for content appropriateness. All of the reviewed stations were deemed to be important and acceptable from a content perspective.

## 5.4 Examiner analyses

Examiner analyses are conducted routinely for each of the 10-minute stations for each PE. For the couplet stations, the examiner analyses are conducted only for the patient interaction component of the couplet stations. For the examiner analyses, the following steps are followed:

### **Step One.**

For each PE and station/component the PE scored, the average across the candidate's station scores is calculated. This average is the PE average for that station or component. Then the average of the PE averages is calculated along with the standard deviation (SD). PEs that scored fewer than 10 candidates on a station are excluded from these analyses as they have observed too few candidates to be compared to other PEs. PEs are flagged as being a "Dove" at 3 SDs if their station or component score is more than three times the station/component SD from the station/component average. PEs are flagged as being a "Hawk" at 3 SDs if their station or component score is less than three times the station/component SD from the station/component average. For example, if the average across PE averages was 72.5 and the SD across PEs was 6.5 and a PE had an average of 50.7 [difference of 21.8, which is greater than 3SDs ( $6.5 \times 3 = 19.5$ )] then he/she is flagged as a "Hawk" at the 3 SD level.

### **Step Two.**

In step two, for each PE flagged in step one, the station distribution (histogram) for the PE is compared to the distribution of station scores from other PEs across the country. This is a visual check to evaluate whether the PE is providing a range of scores that looks somewhat normally distributed (not providing all high or low scores). If a PE's distribution looks reasonable, they are no longer flagged at this step as being either a "Dove" or "Hawk".

### **Step Three.**

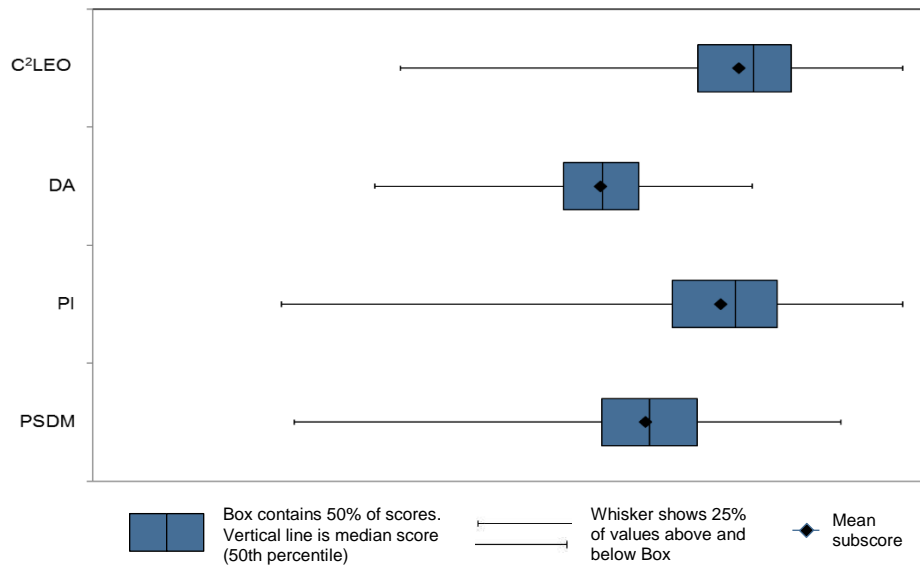
In step three, for each PE flagged in step one and two, the reported score distribution (histogram) for the cohort they scored is compared to the distribution of scale scores from the rest of the candidates across the country. This is a check that the cohort based on all 12 examiners is higher or lower than the values across the country. In this step, we evaluate if a cohort may be higher or lower in ability that may explain a “Dove” or “Hawk” flag in step one. In addition, the average scale scores and pass rate for this PE’s cohort are compared across the country. In this step, we are evaluating whether the cohort of candidates being evaluated may be a weaker or stronger in ability based on all 12 stations. For example, a PE may be flagged as being a “Hawk” in steps one and two, but their scale scores based on all 12 stations may have substantially lower scale scores, indicating a weaker cohort, and thus would not be flagged as a “Hawk” at step three.

One PE in the spring was flagged across all three steps as a “Dove”. This PE was provided a feedback letter on their performance as an examiner, along with the option of a follow-up conversation with the MCCQE Part II Manager.

## **5.5 Domain subscore profiles**

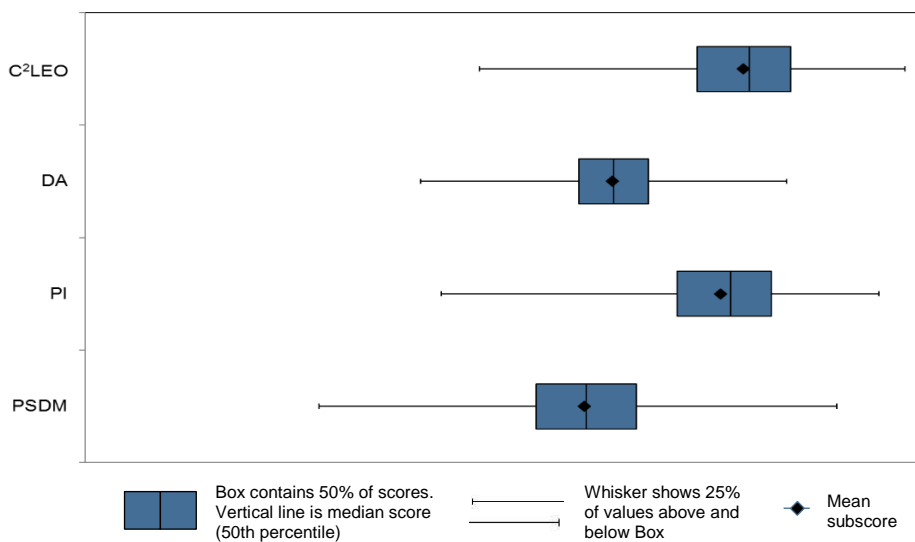
The purpose of the domain subscore profile is to provide diagnostic feedback to candidates by highlighting their relative strengths and weaknesses in four areas. A domain subscore profile is presented in the form of a graph to each candidate in the SFR. The graph shows the domain subscore for each of the four domains and the SEM around the domain subscore. The calculation of the domain subscores for each candidate is outlined in the Domain subscore section within the MCCQE Part II scoring section. This section provides domain subscore profiles for all candidates for the spring and fall MCCQE Part II results. The range of domain subscores for the spring and fall 2017 test forms is shown graphically in Figures 3 and 4, respectively. The boxes for each domain indicate the range for 50 per cent of the candidates’ domain subscores. The vertical line represents the median or 50th percentile domain subscore. The remaining 50 per cent of domain subscores are shown to the right or left of the box as a line (25 per cent to the right and 25 per cent to the left). The mean domain subscores are indicated by the diamond. The note under each figure indicates the naming convention for each domain.





**Figure 3: Domain subscore for spring 2017 MCCQE Part II candidates**

Note: C<sup>2</sup>LEO = Considerations for Cultural-Communication, Legal, Ethical, and Organizational aspects of the Practice of Medicine, DA = Data Acquisition, PI = Patient / Physician Interaction, PSDM = Problem-Solving and Decision Making



**Figure 4: Domain subscore profile for fall 2017 MCCQE Part II candidates**

**Note:** C<sup>2</sup>LEO = Considerations for Cultural-Communication, Legal, Ethical, and Organizational aspects of the Practice of Medicine; DA = Data Acquisition; PI = Patient / Physician Interaction; PSDM = Problem-Solving and Decision Making

## 5.6 Historical pass rates

Historical pass rates since harmonization with the College of Family Physicians of Canada (CFPC) in 2013 are presented in this section. Table 11 shows the pass rates from 2013 to 2017 for four subgroups and the total group of candidates that took the MCCQE Part II or the equivalent CFPC Clinical Skills Exam (CSE) test forms<sup>4</sup>. The four subgroups are CMG-CPG 1st, CMG-CPG Repeat, Other 1st and Other Repeat, where Other is a mix of undergraduate medical education or postgraduate training outside of Canada. Tables 12 and 13 show the pass rates from 2013 to 2017 for the same four subgroups and total group by spring and fall test administration of each year.

**Table 11: 2013 – 2017 pass rates by total reference group**

<b>Year</b>	<b>Candidate group</b>	<b>N total</b>	<b>N pass</b>	<b>Pass rate</b>
<b>2017</b>	<i>Total</i>	4981	4340	87.1
	<i>CMG-CPG 1st</i>	2871	2791	97.2
	<i>CMG-CPG Repeat</i>	265	241	90.9
	<i>Other 1st</i>	1170	870	74.4
	<i>Other Repeat</i>	675	438	64.9
<b>2016</b>	<i>Total</i>	4919	3906	79.4
	<i>CMG-CPG 1st</i>	2969	2719	91.6
	<i>CMG-CPG Repeat</i>	282	226	80.1
	<i>Other 1st</i>	1020	638	62.5
	<i>Other Repeat</i>	648	323	49.8
<b>2015</b>	<i>Total</i>	4806	3611	75.1
	<i>CMG-CPG 1st</i>	2792	2574	92.2
	<i>CMG-CPG Repeat</i>	213	145	68.1
	<i>Other 1st</i>	1162	651	56.0
	<i>Other Repeat</i>	639	241	37.7
<b>2014</b>	<i>Total</i>	4472	3467	77.5
	<i>CMG-CPG 1st</i>	2554	2388	93.5
	<i>CMG-CPG Repeat</i>	149	111	74.5
	<i>Other 1st</i>	1118	694	62.1
	<i>Other Repeat</i>	651	274	42.1
<b>2013</b>	<i>Total</i>	4204	3320	79.0
	<i>CMG-CPG 1st</i>	2476	2361	95.4
	<i>CMG-CPG Repeat</i>	128	102	79.7
	<i>Other 1st</i>	972	578	59.5
	<i>Other Repeat</i>	628	279	44.4

<sup>4</sup> CFPC CSE was discontinued as of spring 2016.

**Table 12: Spring 2013 – 2017 pass rates by total reference group**

<b>Year</b>	<b>Candidate group</b>	<b>N total</b>	<b>N pass</b>	<b>Pass rate</b>
<b>2017</b>	<i>Total</i>	2423	2050	84.6
	<i>CMG-CPG 1st</i>	1163	1139	97.9
	<i>CMG-CPG Repeat</i>	203	188	92.6
	<i>Other 1st</i>	652	466	71.5
	<i>Other Repeat</i>	405	257	63.5
<b>2016</b>	<i>Total</i>	2363	1771	74.9
	<i>CMG-CPG 1st</i>	1248	1120	89.7
	<i>CMG-CPG Repeat</i>	179	144	80.4
	<i>Other 1st</i>	549	317	57.7
	<i>Other Repeat</i>	387	190	49.1
<b>2015</b>	<i>Total</i>	2237	1656	74.0
	<i>CMG-CPG 1st</i>	1158	1109	95.8
	<i>CMG-CPG Repeat</i>	126	88	69.8
	<i>Other 1st</i>	620	350	56.5
	<i>Other Repeat</i>	333	109	32.7
<b>2014</b>	<i>Total</i>	2064	1611	78.1
	<i>CMG-CPG 1st</i>	1082	1047	96.8
	<i>CMG-CPG Repeat</i>	67	46	68.7
	<i>Other 1st</i>	603	386	64.0
	<i>Other Repeat</i>	312	132	42.3
<b>2013</b>	<i>Total</i>	1826	1430	78.3
	<i>CMG-CPG 1st</i>	995	958	96.3
	<i>CMG-CPG Repeat</i>	69	52	75.4
	<i>Other 1st</i>	446	273	61.2
	<i>Other Repeat</i>	316	147	46.5

**Table 13: Fall 2013 – 2017 Pass rates by total reference group**

<b>Year</b>	<b>Candidate group</b>	<b>N total</b>	<b>N pass</b>	<b>Pass rate</b>
<b>2017</b>	<i>Total</i>	2558	2290	89.5
	<i>CMG-CPG 1st</i>	1708	1652	96.7
	<i>CMG-CPG Repeat</i>	62	53	85.5
	<i>Other 1st</i>	518	404	78.0
	<i>Other Repeat</i>	270	181	67.0
<b>2016</b>	<i>Total</i>	2556	2135	83.5
	<i>CMG-CPG 1st</i>	1721	1599	92.9
	<i>CMG-CPG Repeat</i>	103	82	79.6
	<i>Other 1st</i>	471	321	68.2
	<i>Other Repeat</i>	261	133	51.0
<b>2015</b>	<i>Total</i>	2569	1955	76.1
	<i>CMG-CPG 1st</i>	1634	1465	89.7
	<i>CMG-CPG Repeat</i>	87	57	65.5
	<i>Other 1st</i>	542	301	55.5
	<i>Other Repeat</i>	306	132	43.1
<b>2014</b>	<i>Total</i>	2408	1856	77.1
	<i>CMG-CPG 1st</i>	1472	1341	91.1
	<i>CMG-CPG Repeat</i>	82	65	79.3
	<i>Other 1st</i>	515	308	59.8
	<i>Other Repeat</i>	339	142	41.9
<b>2013</b>	<i>Total</i>	2378	1890	79.5
	<i>CMG-CPG 1st</i>	1481	1403	94.7
	<i>CMG-CPG Repeat</i>	59	50	84.7
	<i>Other 1st</i>	526	305	58.0
	<i>Other Repeat</i>	312	132	42.3

## 6. References

---

- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Sage Publications Inc.: Thousand Oaks, CA.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart & Winston, Inc.: Orlando, FL.
- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65-110). Westport, CT: Praeger Publishers.
- Kane, M.T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp.17-64). Westport, CT: Praeger Publishers.
- Kane, M.T. (2013). Validating the interpretations and uses of test scores. *Journal of educational measurement*, 50(1), 1-73.
- Kolen, M.J., & Brennan, R.L., (2004). *Test equating, scaling, and linking: Methods and practice* (2nd ed.). New York, NY: Springer Science Business Meeting.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of educational measurement*, 32(2), 179-197.  
doi: 10.1111/j.1745-3984.1995.tb00462.x
- Medical Council of Canada. (2014). *Blueprint project: Qualifying examinations blueprint and content specifications*. MCC: Ottawa, Canada. Retrieved from:  
[mcc.ca/wp-content/uploads/Blueprint-Report.pdf](http://mcc.ca/wp-content/uploads/Blueprint-Report.pdf)
- Medical Council of Canada. (2015). *Technical report on the standard setting exercise for the Medical Council of Canada Qualifying Examination Part II*. MCC: Ottawa, Canada. Retrieved from:  
[mcc.ca/wp-content/uploads/MCCQE-Part-II-Standard-Setting-Report-July2015.pdf](http://mcc.ca/wp-content/uploads/MCCQE-Part-II-Standard-Setting-Report-July2015.pdf)

## Appendix A: Quality control – MCCQE Part II results

---

### **Pre-examination**

<b>1.</b>	Design sheets, set up data exports and activate sheets for all stations in TeleForm®; print, bubble, hand score and scan 10 test sheets per station, per organization, per language and perform quality assurance (QA) checks of raw data in CSV files	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>2.</b>	Create answer keys, compare against exam sheets, case, scoring rules and adjust as necessary; create and QA post-encounter probe (PEP) answer keys and other docs for PEP marking	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>3.</b>	Enter answer key into SQL database and download answer key to SAS	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>4.</b>	Update the SAS code as necessary (new question type, etc.)	
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>5.</b>	Import CSV raw data into scoring application and run scoring on fake candidates, run SAS scoring in parallel and verify matching results in SAS and scoring application	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
	Senior Psychometrician (Overview of analyses)	<i>(Initials &amp; sign-off date)</i>
<b>6.</b>	Test candidate bar code labels	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>7.</b>	Test examiner bar code labels	
	Administration Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>8.</b>	Prepare a file with demographic information about candidates including creating candidate groups (candidate info)	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
	Senior Psychometrician (Overview of analyses)	<i>(Initials &amp; sign-off date)</i>

### **Post-examination (Pre-CEC)**

These tasks **MUST** be completed before the CEC meets.

<b>9.</b>	On-site scanning, receiving and uploading data files from sites and load files into TeleForm®	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>10.</b>	Verify	
	CSA team	<i>(Initials &amp; sign-off date)</i>
<b>11.</b>	PEP marking, scan PEP sheets, upload files to TeleForm® and verify	
	Production Coordinator/Test Development Officer, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>12.</b>	Import CSVs into scoring application and correct until error reports are clean	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>13.</b>	Run SAS scoring in parallel with scoring in SQL; verify matching results in SAS and scoring application	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>14.</b>	Run pull lists and QA against SAS	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Production Coordinator, CSA	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>15.</b>	Pull sheets, hand score, enter into hand scoring application and update CSV files	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Production Coordinator, CSA	<i>(Initials &amp; sign-off date)</i>
<b>16.</b>	PEP QA	
	Production Coordinator/Test Development Officer, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>17.</b>	Run preliminary statistical analyses	
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>18.</b>	Review statistical information	
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>

<b>19.</b>	Present preliminary statistics to the Manager, MCCQE Part II and Associate Director, Evaluation Bureau	
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>
<b>20.</b>	If the statistical analyses indicate an unusual pattern, call a meeting with the Manager, MCCQE Part II, Associate Director, Evaluation Bureau, Senior Research Psychometrician and Analysts (SQL and SAS)	
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>
	<p>Note unusual pattern:  * Add any supporting documentation to the folder.  * Additional notes or comments can be added to the last page of this document.</p>	
<b>21.</b>	Hand score and data entry changes from PEP QA into hand score application; run error reports and correct errors to hand score data	
	Temp staff	<i>(Initials &amp; sign-off date)</i>
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>22.</b>	Enter data changes in CSV files from PEP QA, QA data entry and re-run scoring	
	National Site Coordinator/other MCCQE Part II staff	<i>(Initials &amp; sign-off date)</i>
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>23.</b>	Run SAS scoring in parallel with scoring application; verify matching results in SAS and scoring application	
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>24.</b>	Other QA reports and corrections (i.e., candidates with zero scores on components, examiner analyses, missing data, etc.)	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>25.</b>	Examiner feedback: data entry and editing	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>26.</b>	Incident reports: data entry, classification, and follow-up	
	Administration Coordinator (A)	<i>(Initials &amp; sign-off date)</i>
	Administration Coordinator (B)	<i>(Initials &amp; sign-off date)</i>
<b>27.</b>	Patient safety and lapses in professional behaviour	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
<b>28.</b>	Special case investigations	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>



	Manager, Data Analyses (overview of Analysts)	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician (overview)	<i>(Initials &amp; sign-off date)</i>
<b>29.</b>	Super borderline review	
	Production Coordinator and team, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>30.</b>	Ad-hoc investigations	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>31.</b>	CEC teleconference	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>32.</b>	Receive results from CFPC	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>33.</b>	Prepare reports to CEC (special cases)	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>34.</b>	Update CSV files and re-score (if necessary)	
	QEII Production Coordinator	<i>(Initials &amp; sign-off date)</i>
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>35.</b>	Re-run all quality assurance steps	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician (final review)	<i>(Initials &amp; sign-off date)</i>
<b>36.</b>	Prepare CEC report	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>
<b>37.</b>	Dry run and approval of the CEC report	
	Associate Director, Evaluation Bureau	<i>(Initials &amp; sign-off date)</i>
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>
<b>38.</b>	Test import into physiciansapply.ca (staging)	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>39.</b>	QA SoR and SFR from staging environment	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>

**CEC**

<b>40.</b>	<b>Present CEC report and obtain approval of results</b>	
	Senior Research Psychometrician (review item)	<i>(Initials &amp; sign-off date)</i>
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>41.</b>	<b>Note/document decisions by CEC</b>	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
<b>42.</b>	<b>Update SQL and SAS with decisions by CEC</b>	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician (review item)	<i>(Initials &amp; sign-off date)</i>
<b>43.</b>	<b>Export physiciansapply.ca data from scoring application to Access database</b>	
	Analyst (SQL)	<i>(Initials &amp; sign-off date)</i>
<b>44.</b>	<b>QA Access</b>	
	Analyst (SAS)	<i>(Initials &amp; sign-off date)</i>
<b>45.</b>	<b>Import to physiciansapply.ca (production), validate and approve</b>	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Associate Director, Evaluation Bureau	<i>(Initials &amp; sign-off date)</i>
	Director, Evaluation Bureau	<i>(Initials &amp; sign-off date)</i>
<b>46.</b>	<b>QA production SoR and SFR</b>	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>
	Senior Research Psychometrician	<i>(Initials &amp; sign-off date)</i>
<b>47.</b>	<b>Updating the CEC report with the CEC decisions (last slide)</b>	
	Manager, MCCQE Part II	<i>(Initials &amp; sign-off date)</i>

## Appendix B: Patient interaction rating scales

---

### 1. Initiation of interview [ONE bubble only]

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lack of introduction	Minimal acknowledgement of patient	Borderline unsatisfactory; acknowledges patient, introduces self	Borderline satisfactory; acknowledges patient, introduces self	Acknowledges patient; moderately at ease and attentive	Attentive to patient; introduces self; at ease, personable

### 2. Listening skills [ONE bubble only]

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Interrupts inappropriately, ignores patient's answers	Impatient	Borderline unsatisfactory; somewhat attentive	Borderline satisfactory;	Interrupts inappropriately, ignores patient's answers	Impatient

### 3. Questioning skills [ONE bubble only]

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Awkward, exclusive use of closed-ended or leading questions and jargon	Somewhat awkward; inappropriate terms; minimal use of open-ended questions	Borderline unsatisfactory; moderately at ease; appropriate language; uses different types of questions	Borderline satisfactory; moderately at ease; appropriate language; uses different types of questions	At ease; clear questions; appropriate use of open and closed-ended questions	Confident; skillful questioning

### 4. Organization of interview [ONE bubble only]

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scattered, shot-gun approach	Minimally organized	Borderline unsatisfactory; somewhat logical flow	Borderline satisfactory; logical flow	Logical flow with sense of purpose	Purposeful, integrated handling of encounter

### 5. Rapport with person [ONE bubble only]

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Condescending, offensive, judgmental	Minimal courtesies only	Borderline unsatisfactory	Borderline satisfactory	Polite and interested	Warm, empathic

**6. Information giving [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No attempt or inappropriate attempt to give information (e.g., not truthful)	Awkward and / or incomplete attempts to give information	Borderline unsatisfactory; somewhat at ease, attempts to give information	Borderline satisfactory; somewhat at ease, attempts to give information	Gives information easily; somewhat attentive to patient's understanding	Confident and skillful at giving information; attentive to patient's understanding (e.g., truthful)

**7. Professional behaviour with patient [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Offensive or aggressive; frank exhibition of unprofessional conduct	Negative attitude toward patient	Borderline unsatisfactory; does not truly instill confidence	Borderline satisfactory; manner inoffensive, but does not necessarily instill confidence	Attempts professional manner with some success	Overall demeanour of a professional; caring, listens, communicates effectively

**8. Professional behaviour with colleague [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Offensive or aggressive; frank exhibition of unprofessional conduct	Negative attitude	Borderline unsatisfactory; not truly respectful	Borderline satisfactory; moderately respectful	Demonstrates professional manner with some success	Overall demeanour of a professional; listens, communicates effectively

**9. Ethical conduct [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Markedly inappropriate or awkward handling of ethical issues	No consideration of ethical issues	Borderline unsatisfactory; minimal consideration of ethical issues	Borderline satisfactory; minimal and appropriate consideration of ethical issues	Responds satisfactorily to ethical issues	Considers ethical issues with care and effectiveness, and responds to them

**10. Organization of physical examination [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exam not attempted or scattered; patient moved unnecessarily	Minimally organized	Borderline unsatisfactory; somewhat logical flow	Borderline satisfactory; logical flow	Logical flow with sense of purpose	Purposeful, integrated handling of examination

**11. Attention given to patient's physical comfort [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exam not attempted, or inattentive to patient's comfort or dignity (e.g., no draping, causes pain unnecessarily)	Causes some unnecessary discomfort or embarrassment	Borderline unsatisfactory in attending to patient's comfort and needs	Borderline satisfactory in attending to patient's comfort and needs	Mostly attentive to patient's comfort and dignity	Consistently attentive to patient's comfort and dignity

**12. Demonstration of technical skills [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Exam not attempted or maneuvers cannot provide reliable / useful information	Maneuvers too rushed or clumsy; unlikely to provide reliable / useful information	Borderline unsatisfactory; some skill, but minimal likelihood of reliable / useful findings	Borderline satisfactory; some skill, some reliable / useful findings likely	Consistent skill; maneuvers likely to provide reliable / useful information	Consistent skill; maneuvers performed will elicit reliable / useful information

**13. Relationship to the patient [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Introduction absent or inappropriate; no consent; awkward; uses jargon; no acknowledgment of patient	Minimal interaction and / or minimal acknowledgment of patient	Borderline unsatisfactory in approach to patient	Borderline satisfactory in approach to patient	Moderately clear and understandable; acknowledges patient; moderately at ease with patient	Clear, concise instructions; elicits consent to physical examination; at ease with patient

**14. Overall organization of patient encounter [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No logical flow; scattered, inattentive to patient's agenda	Counsels patient before taking history or doing physical	Minimal organization; scattered approach	Appropriate approach to patient	Skillful approach to patient	Skillful, professional approach to patient and effective use of time

**15. Compliance optimization [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Candidate's approach may negatively affect compliance	Patient's compliance unlikely to be optimized	Borderline unsatisfactory; weak attempt to encourage patient's compliance	Borderline satisfactory; candidate's approach may positively affect patient's compliance	Candidate's approach encourages patient's compliance	Candidate's approach highly likely to optimize patient's compliance


**16. Facilitation of informed decision making [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No attempt or inappropriate attempt at information sharing (e.g., deception, slanting of facts, incorrect information)	Incomplete and / or biased information; overuses jargon; does not ensure understanding of issues	Attempts to share information; omits some critical facts; uses some jargon; attempts to ensure understanding	Gives some information on most important facts; may use jargon; attempts to ensure understanding	Gives clear information; supports patient decision making (e.g., alternatives, risks/benefits); appropriate language; ensures understanding	Organized; optimizes patient decision making; significant effort to make information relevant; clear language; attentive to patient understanding

**17. Responds to question [ONE bubble only]**

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ignores question or is rude, dismissive	Poor answer or explanation	Borderline unsatisfactory answer or explanation	Borderline satisfactory answer or explanation	Clear answer or explanation; is empathetic during interaction	Clear answer or explanation; empathetic; asks if understands or needs more information

## Appendix C: Example SOR and SFR



MEDICAL COUNCIL OF CANADA LE CONSEIL MÉDICAL DU CANADA

1021 Thomas Spratt Place  
1021, place Thomas Spratt  
Ottawa, ON  
Canada K1G 5L5  
613-521-6012

**MEDICAL COUNCIL OF CANADA  
QUALIFYING EXAMINATION PART II  
STATEMENT OF RESULTS**

---

**November 28, 2017**


<b>Candidate name:</b>	XXXXXXXX, XXXXXX	<b>Your final result:</b>	Pass
<b>MCC candidate code:</b>	XXXXXXXXXX	<b>Your total score:</b>	598
<b>Examination session:</b>	MCCQE Part II Fall 2017	<b>Score required to pass:</b>	509

On behalf of the Central Examination Committee, I am writing to inform you of your final result on the Medical Council of Canada Qualifying Examination (MCCQE) Part II that took place during the above-mentioned session.


Your total score, which represents your overall performance, is reported as a scaled score ranging from 50 to 950. Your final result (e.g., pass/fail) is based on your total score relative to the score required to pass. Additional information, including the mean and standard deviation, is available from the [MCCQE Part II Scoring web page](#).

Supplemental feedback on your examination performance is reported to you in a separate document within your [physiciansapply.ca](#) account.

Please accept my best wishes for future success.



M. Ian Bowmer, MD CM, FRCPC  
Executive Director and Registrar  
Medical Council of Canada



---

mcc.ca  
physiciansapply.ca  
inscriptionmed.ca



1021 Thomas Spratt Place  
1021, place Thomas Spratt  
Ottawa, ON  
Canada K1G 5L5  
613-521-6012

## SUPPLEMENTAL FEEDBACK REPORT

**Candidate name:** XXXXXXXX, XXXXXX

**MCC candidate code:** XXXXXXXXXX

**Your total score:** 598

**Examination:** MCCQE Part II

**Examination session:** Fall 2017

The purpose of this report is to provide you with supplemental information on your relative strengths and weaknesses, based on your performance across the different domains that were assessed by the test form of the Medical Council of Canada Qualifying Examination (MCCQE) Part II that was administered to you.

Figure 1 displays your performance measured across stations in the following four domains: C<sup>2</sup>LEO (Considerations for Cultural-Communication, Legal, Ethical, and Organizational aspects of the Practice of Medicine), Data Acquisition, Patient / Physician Interaction, and Problem Solving and Decision Making. Each domain is sampled a number of times, with some being measured across all stations and others across a smaller number of stations.

- The C<sup>2</sup>LEO subscore is calculated using checklist items, oral questions, rating scales, and/or written questions across the stations that measure the C<sup>2</sup>LEO domain.
- The Data Acquisition subscore is calculated using items within stations that pertain to history taking and physical examination.
- The Patient / Physician Interaction subscore is calculated using rating scale items within most of the stations.
- The Problem Solving and Decision Making subscore is calculated using oral question items, relevant items within the management stations, as well as items from the written stations.

To help you better understand your performance, your subscore for each domain is shown along with the mean score of candidates from this examination session who were first-time takers of the MCCQE Part II and who passed. The standard error of measurement (SEM) associated with each of your subscores represents the expected variation in your subscore if you were to take this examination again with a different set of stations covering the same or similar domains. Small differences in subscores or overlap between SEMs are indicative that performance in those domains was relatively similar. Likewise, overlap between the SEM for a domain subscore and the mean score of first-time takers who passed, within a given domain, signifies that performance is similar to the mean.

**Report date:** 2017-11-28

1/2

mcc.ca  
physiciansapply.ca  
inscriptionmed.ca

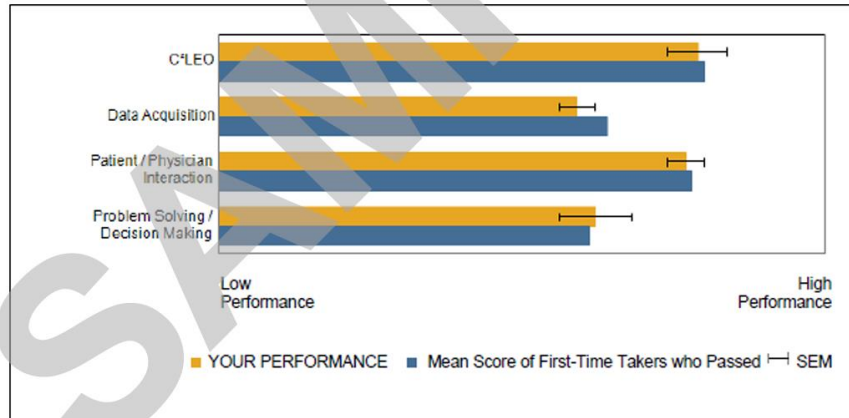


### SUPPLEMENTAL FEEDBACK REPORT

It is important to note that the subscores are based on significantly less data than the total score and that these do not have the same level of precision as the total score. If you have failed the examination and wish to retake it, preparation for all domains is important; otherwise you could improve some subscores and inadvertently lower others. Please also note that your subscores cannot be directly compared to your total score in that they are calculated differently and as such the pass score cannot be translated to Figure 1.

For more information, please visit the [MCCQE Part II Scoring web page](#).

Figure 1. MCCQE Part II Score Profile



Report date: 2017-11-28

2/2

Xxxxxxxxx, Xxxxxx / XXXXXXXXXXXX

mcc.ca  
physiciansapply.ca  
inscriptionmed.ca