Medical
Council
of Canada
Qualifying
Examination
(MCCQE)
Part II

MCCQE PART II 2019
ANNUAL
TECHNICAL
REPORT

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

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 2019. 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 in 2019 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 candidate's core abilities to apply medical knowledge, demonstrate clinical skills, develop investigational and therapeutic clinical plans, as well as demonstrate professional behaviours and attitudes at a level expected of a physician in independent practice in Canada.

The exam is a 12-station objective structured clinical examination (OSCE) that is based on the Blueprint (Medical Council of Canada, 2014), and assesses the candidate's performance across two broad categories:

1. Dimensions of care, covering the spectrum of medical care, and
2. Physician activities, reflecting a physician's scope of practice and behaviours.

The exam consists of a series of 10 stations that count towards the candidate's total score and two pilot stations that do not count. For each station, a brief, written statement introduces a clinical problem and directs the candidate to appropriately examine a standardized participant (SP), who simulates a patient, 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 interact with simulated health professionals or other simulated roles (i.e., the parent of a child patient). 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 healthcare.

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, in May and October. It is scored by physician examiners (PEs) or non-physician examiners (NPEs) 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 blueprint and specifications

The Blueprint as outlined in Table 1 (Medical Council of Canada, 2014) was developed for the MCCQE Part I and the MCCQE Part II and approved by Council in September 2014. The blueprint is two-dimensional and designed to assess candidates' performance across two broad categories:

1. Dimensions of care, covering the spectrum of medical care, and
2. Physician activities, reflecting a physician's scope of practice and behaviours.

Table 1: Exam blueprint for the MCCQE Part II

	Dimensions of care			
	Health Promotion & Illness Prevention	Acute	Chronic	Psychosocial Aspects
Assessment/ Diagnosis				
Management				
Communication				
Professional Behaviours				

Physician activities

Each category has four domains, and each is assigned a specific content weighting on the exam:

Dimensions of Care
Reflects the focus of care for the patient, family, community and/or population

Health Promotion and Illness Prevention
The process of enabling people to increase control over their health and its determinants, and thereby improve their health. Illness prevention covers measures not only to prevent the occurrence of illness, such as risk factor reduction, but also to arrest its progress and reduce its consequences once established. This includes, but is not limited to screening, periodic health exam, health maintenance, patient education and advocacy, and community and population health.

Acute

Brief episode of illness within the time span defined by initial presentation through to transition of care. This dimension includes but is not limited to urgent, emergent, and life-threatening conditions, new conditions, and exacerbation of underlying conditions.

Chronic

Illness of long duration that includes but is not limited to illnesses with slow progression.

Psychosocial Aspects

Presentations rooted in the social and psychological determinants of health and how these can impact on wellbeing or illness. The determinants include but are not limited to life challenges, income, culture, and the impact of the patient's social and physical environment.

Physician Activities

Reflects the scope of practice and behaviours of a physician practising in Canada

Assessment/Diagnosis

Exploration of illness and disease using clinical judgment to gather, interpret and synthesize relevant information that includes but is not limited to history taking, physical examination and investigation.

Management

Process that includes but is not limited to generating, planning, organizing safe and effective care in collaboration with patients, families, communities, populations, and other professionals (e.g., finding common ground, agreeing on problems and goals of care, time and resource management, roles to arrive at mutual decisions for treatment, working in teams).

Communication

Interactions with patients, families, caregivers, other professionals, communities and populations. Elements include but are not limited to relationship development, intraprofessional and interprofessional collaborative care, education, verbal communication (e.g. using the patient-centered interview and active listening), non-verbal and written communication, obtaining informed consent, and disclosure of patient safety incidents.

Professional Behaviours

Attitudes, knowledge, and skills relating to clinical and/or medical administrative competence, communication, ethics, as well as societal and legal duties. The wise application of these behaviours demonstrates a commitment to excellence, respect, integrity, empathy, accountability and altruism within the Canadian health-care system. Professional behaviours also include but are not limited to self-awareness, reflection, life-long learning,

leadership, scholarly habits and physician health for sustainable practice. (Medical Council of Canada, 2014, p.7)

The exam blueprint was used to inform exam specifications for the MCCQE Part II by assigning a weight to each domain and applying additional constraints.

Table 2 provides the exam specifications, which includes each category and domain and its associated weight for the MCCQE Part II.

Table 2: Exam blueprint for the MCCQE Part II with weighted domains

Dimensions of care					
Physician activities	Health Promotion & Illness Prevention	Acute	Chronic	Psychosocial Aspects	Row %
Assessment/ Diagnosis					25±5
Management					25±5
Communication					25±5
Professional Behaviours					25±5
Column %	20±5	35±5	25±5	20±5	100

In addition, the exam specifications include the following constraints for each exam form:

- At a minimum, each test form must have one child patient, one elderly patient and four adult patients
- At a maximum, each test form must have no more than two cases per body system (e.g., cardiovascular, respiratory, etc.)
- At a minimum, each test form must be comprised of 40 per cent male patients and 40 per cent female patients
- At a minimum, each test form must be comprised of 20 per cent multi-morbid / complex stations
- Each test form must have two to three cases with a physical examination
- Each test form must have one emergent care case

- Each test form must have one Health Professional interaction
- At a maximum, each test form must have five inpatient cases, at least one in a rural setting
- There should be a representative sampling of special populations, included but not limited to immigrant, ability to access care, disabled, First Nation, Inuit, and Métis populations; end of life patients, refugees, inner city poor, the addicted and the homeless.

Creating exam forms that meet exam specifications leads to forms measuring similar content. This (along with linking, described in Section 4.3.1) allows candidates' exam scores to be compared fairly from one exam form to the next.

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, as described below, 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 blueprint for competent physicians in independent practice in Canada.

Case authors create stations that evaluate candidates' ability to assess and diagnose, manage, communicate and behave professionally in a range of clinical situations commonly encountered by physicians.

Case authors first develop the instructions given to candidates prior to entering an OSCE station. The instructions include 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.

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. Candidates attempting the MCCQE Part II completed an exam form composed of eight 14-minute encounter stations on Day 1 (Saturday), and four paired stations on Day 2 (Sunday). Paired stations consist of an encounter with a SP, either preceded by a task such as reading a chart or followed by presenting a report of the patient to the examiner. The paired 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 obtain a history, conduct a physical exam, or assess and manage the patient's issues.

Standardized administration, examiner/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.

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 exam 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 specifications. 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, specifically 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 resulting exam specification characteristics for the MCCQE Part II forms that were administered in 2019. The "Target" column specifies the target percentages for each exam form for each dimension of care and physician activity (as shown in Table 1). Any deviation from the targets are approved by both the MCCQE Part II Test Committee and the Central Examination Committee. One exam form was administered in May and one in October 2019.

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

Dimensions of care	Target	May	October
Health Promotion and Illness Prevention	20% ± 5%	25%	18%
Acute	35% ± 5%	41%	44%
Chronic	25% ± 5%	20%	21%
Psychosocial Aspects	20% ± 5%	14%	17%
Physician activities	Target	May	October
Assessment / Diagnosis	25% ± 5%	30%	33%
Management	25% ± 5%	23%	19%
Communication	25% ± 5%	21%	23%
Professional Behaviours	25% ± 5%	26%	25%

2.6 Exam scoring

Each OSCE station item is assigned to one dimension of care (health promotion and illness prevention, acute, chronic, psychosocial aspects) and one physician activity (assessment / diagnosis, management, communication, professional behaviours). Examiners use checklists, oral question items, and rating scales[1] to score the candidate's proficiency on each station. Different examiners evaluate candidates in each of the stations, and one set of station scores per candidate is collected. The ratings provided by each examiner are used to calculate all scores.

In almost all stations, the examiner 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.

---

[1] Rating scales are four-level Likert items with scores ranging from zero to three. 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 PGT or osteopathic postgraduate training (PGY-2+); or completed 12 months of PGT on, or before the deadline of June 30 for the May exam of the same calendar year or December 31 for the October exam of the same calendar year (PGY-1). The MCC implemented capacity limits to registration for the May and October administrations beginning in October 2015 due to the demand in particular sessions exceeding the available number of test spots and the concern that increased demand would affect the quality of the examination administration. Additionally, implementing a 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 2019 exam administrations, candidates that met the eligibility requirements were invited to add their names to the MCCQE Part II pre-application list through their physiciansapply.ca account during specific pre-application periods. When each pre-application period ended, candidates were randomly selected from the pre-application list based on available exam capacity.

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, an explanation of scoring and results, application information as well as sample stations and an online OSCE orientation. Candidates must also participate in a mandatory orientation given 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 candidates will be assessed by the examiners and when an examiner may intervene

- Information on medical equipment and props available in the room
- A description of how candidates should interact with SPs
- An overview of the extended match written stations
- Instructions on exam security and how to ask for assistance
- MCC exam security video
- 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 2019, 4,589 candidates participated in the MCCQE Part II. The exam was 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), examiners, exam day staff, caterers, etc.) perform to standards as outlined in the Site Administration Manual.

MCC personnel oversees site staff on exam days across the country in person, by telephone, and via electronic communication 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 examiner recruitment and enhance collaboration between sites.

3.5 Standardized Participants

SPs are healthy individuals (or individuals with chronic stable findings) who are trained to present a patient's signs and/or symptoms reliably, consistently and realistically.

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 training materials provided by the MCC. Training support is provided to the SP trainers by MCC staff, primarily by the training officer.

All SPs take part in a dry run prior to exam day with the CE playing the candidate's role to ensure

that the SPs are ready to perform their roles. An SP who is deemed as not being ready is re-trained or replaced.

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

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 MCCs clinical skills assessments
- Providing leadership and support to senior site staff for planning and teamwork
- With assistance of the CE (see below), 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 (QA), 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), examiners, 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. Before the exam, the CE assists with SP dry runs, participates in some of the SP training sessions and assists in examiner recruitment and staff training as needed. On the day of the exam the CE responds to examiner questions, assists candidates as needed, may deliver the examiner orientation, and supports the administrative team as needed.

3.8 Examiner recruitment and training

Examiners are physicians, nurses, other health professionals or highly trained SPs. Of the 10 stations that count towards the candidate's total score, no more than two stations may be examined by an NPE.

Table 4 presents the requirements used to recruit PEs.

Table 4: PE recruitment requirements

REQUIREMENTS FOR PE RECRUITMENT
1. PEs must be registered and in good standing with a Medical Regulatory Authority (MRA) in Canada.
2. PEs can be semi-retired or retired, but they must have an active licence with an MRA in Canada.
3. PEs must agree to adhere to the MCC's Code of Business Conduct (the Code): • The Code addresses, but is not limited to, the application of the Code, legal and ethical standards, standards of conduct and compliance with the Code • Any breach of the Code could lead to legal and/or disciplinary measures against the responsible persons, including without limitation their removal from office, their expulsion from committees or the termination of the applicable contract
4. PEs must have the ability and stamina to complete the task (e.g., uncorrected hearing loss is a serious handicap).
5. PEs must be practising or have practised in Canada within the last five years.
6. PEs must have the Licentiate of the Medical Council of Canada (LMCC) and must provide their LMCC registration number. Examiners must be two years post-LMCC.
7. PEs must have at least two years of independent practice in Canada (exception: for residents, see number 9).

REQUIREMENTS FOR PE RECRUITMENT
8. PEs do not need to be faculty members if all other criteria are met AND they meet the following criterion: • They must have recent experience supervising clerks and residents (within the last two years) (e.g., a physician who has worked for a university as an OSCE examiner or for other similar OSCEs)
9. PEs can be residents, but they must be PGY5 level or higher, or they can be fellows, OR they can have CCFP certification. • **Note:** Residents must not exceed 20 per cent of the examiner cohort assigned to a given session

10. Physicians who do not have their LMCC will be accepted as examiners under the following conditions:
• Non-licentiate PEs must be faculty members (e.g., faculty lecturer, assistant professor, associate professor or professor)
 	
AND
 	
• Non-licentiate PEs must be certified by one of the following organizations and must provide their certification number:
    ○ Royal College of Physicians and Surgeons of Canada (RCPSC)
    ○ Collège des médecins du Québec (CMQ)
    ○ College of Family Physicians of Canada (CFPC)
 	
AND
 	
• Non-licentiate PEs must sign a waiver indicating that they have no intention of taking the MCCQE Part II
 	
Note: Non-licentiate PEs must not exceed 50 per cent of the examiner cohort assigned to a given session.

11. Any potential PEs who do not meet the above guidelines must be pre-approved by the MCCQE Part II Exam Manager. Documentation of qualifications may be requested.

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

4. MCCQE PART II SCORING

This section describes QA 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, QA steps are performed as outlined below.

Examiners complete score sheets for each candidate observed in their OSCE stations. These sheets 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 example, a non-scannable candidate barcode label, an examiner'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 are imported electronically into a scoring application and preliminary scores are calculated. This preliminary set of scores are used to generate a list of all candidates whose score is within a 3-point range around 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 maximum number of correct answers
- Notes by examiners for oral or SP questions
- Confirmation of missing data for oral questions or rating scales
- Flags for lapses in patient safety or professional behaviours
- Verification of raw score points

Any differences are corrected in the electronic data files to reflect the paper score sheets. The updated electronic files are then re-imported into the scoring application that is used to calculate scores for all candidates. All scores are calculated independently 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 (described below). Once the CEC has approved the results, they are imported into physiciansapply.ca and released to candidates.

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 (for example, illness, fire alarm or a SP misportrayal), the CEC may decide to remove a station 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 "No Standing" does not count towards a candidate's number of attempts.

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. A "Denied Standing" counts as an attempt towards a candidate's total number of attempts. Additionally, 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 one week from results being released to candidates, the MCC issues a Statement of Results (SOR) and Supplemental Information Report (SIR) to each candidate through their physiciansapply.ca account (samples from May 2019 administration are shown in Appendix C). The SOR includes the candidate's final result and total score, as well as the pass score. The SIR includes the candidate's final result and total score and additional information in graphic form about the candidate's domain subscores and comparative information.

The total score is reported on a standard-score scale ranging from 50 to 250. In contrast, the score profile in Figure 1 of the SIR displays a candidate's domain subscores in terms of a percentage. As a result, total scores cannot be compared to domain subscores in the SIR as they are reported on different scales. Additionally, it is important to note that, because subscores have fewer items than total scores, subscores have less measurement precision. Subscores are provided to 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

The scale score is the candidate's total score reported on a scale that ranges from 50 to 250 (as opposed to a candidate's total raw score that is on a percentage metric). Deriving the scale score for the 2019 MCCQE Part II involves three steps.

Step 1: Calculate total raw scores

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

$$\text{station score} = \frac{\sum_i \text{score}_i}{\sum \text{max.} \_\text{score}_i} * 100$$

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 raw score for each candidate using the following formula:

$$\text{total raw score} = (\text{sum of 10 station scores})/10$$

Missing data occurs when the examiner does not provide a rating for one or more of the oral questions or rating scales for a given candidate on the score sheet. 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. When oral questions or rating scales are not scored by the examiner, the station score is based on the items that are provided by the examiner. 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 and the station score based on accounting for missing data is used instead.

Step 2: Link MCCQE Part II scores to base form – October 2018

For each MCCQE Part II administration, a different exam 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 seven 14-minute stations and three paired stations that best meet exam specifications. While good for security reasons,

one possible adverse effect of having different exam forms with different stations across administrations is that one exam form may be more difficult than the other. The process of linking total scores statistically takes into account differences in exam form difficulty and adjusts total scores so that all scores are on the same scale. Linking is also a way of applying the same pass score to candidates who take different exam forms.

One method of linking exam forms is to have a subset of the content appear identically across exam 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 exam form and the difficulty of the exam forms into account.

For the MCCQE Part II 2019 exam forms, the anchor set was based on three stations. A reference group of Canadian medical graduate and Canadian postgraduate first-time test takers was used for all linking calculations. The linking calculations from this reference group is applied to all candidates to calculate each candidate's linked score. This linked score is then transformed as described in step 3 below. In all linking steps, the Tucker 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 total raw scores to scale scores

Once the second step is complete, the linked scores are transformed to scale scores ranging from 50 to 250 for reporting purposes.

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

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

Where ScaleScore_X is defined as the linear function to calculate the scale score for candidate X, where the slope is equal to 2.23 based on the transformation of the October 2018 MCCQE Part II, where the intercept is equal to 17.87 based on the transformation of the October 2018 MCCQE Part II, and the LinkedScore_X is the linked score for candidate X.

All scale scores are rounded to a whole number between 50 and 250. The reported scale scores as seen by candidates are these rounded values. For example, a passing candidate with a linked score of 81.25 would have a scale score of 199:

$$\text{ScaleScore}_X = (2.23) * (81.25) + (17.87) = 199.06 \text{ rounded to } 199$$

A failing candidate with a linked score of 42.51 would result in a scale score of 113:

$$\text{ScaleScore}_X = (2.23) * (42.51) + (17.87) = 112.67 \text{ rounded to } 113$$

4.3.2 Pass/fail status

The The MCC completed a rigorous standard-setting exercise using the borderline group method and a panel of 20 physicians from across the country that represented faculties of medicine, different specialties, and differing years of experience supervising residents. The October 2018 MCCQE Part II exam form was used to establish the pass score. The panelists reviewed stations, content and score information and provided judgments for establishing the recommended 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, 2019). The pass score of 138 was recommended by the panel of physicians for October 2018 and approved by the CEC in January 2019. This pass score was used to assign a pass/fail status to each candidate of the 2019 administrations. A score of 138 or greater is deemed a pass.

4.3.3 Domain subscores

Domain subscore calculations are used to create the figure in the candidates' SIRs. 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, 9 and 1, 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 x 100 or 70.0. For the MCCQE Part II, there are eight domain subscores that are presented to candidates in their SIRs:

Dimensions of Care

- Health Promotion and Illness Prevention
- Acute Care
- Chronic Care
- Psychosocial Aspects

Physician Activities

- Assessment and Diagnosis
- Management
- Communication
- Professional Behaviours

As a reminder, domain subscores should not be compared to total scores as they are reported on different scales and because they have fewer items than total scores they have less measurement precision than total scores. Domain subscores are intended to provide 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 for the May 2019 administration and December for the October 2019 administration. In total, 4,589 candidates participated in the MCCQE Part II administered in May and October 2019.

Table 5 shows the number of candidates for the May and October 2019 administrations by candidate group (for example, Canadian Medical Graduates (CMG), first-time test takers (1st), etc.), gender and examination language. When referring to candidate's medical graduation, we separate these candidates in two categories based on where candidates obtained their medical degree, either within Canada or internationally. When referring to candidate's postgraduate training, we separate these candidates into two categories based on where they have completed their residency, either within Canada or internationally. The reference group used for linking and for secondary analyses includes CMGs, Canadian postgraduates, and first-time test-takers. Candidate groups shown in Table 5 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 5: Number and percentage of candidates for the MCCQE Part II by PGT group

PGT Group	Subgroups	May N	May %	October [a] N	October [a] %	Total N	Total %
PGY-1	CMG-CPG 1st	740	90.0	6	85.7	746	90.0
	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	1	0.1	0	0.0	1	0.1
	IMG-IPG repeaters	0	0.0	0	0.0	0	0.0
	IMG-CPG 1st	80	9.7	1	14.3	81	9.8
	IMG-CPG repeaters	1	0.1	0	0.0	1	0.1
	Gender	**N**	**%**	**N**	**%**	**N**	**%**
	Female	453	55.1	4	57.1	457	55.1
	Male	369	44.9	3	42.9	372	44.9
	Language	**N**	**%**	**N**	**%**	**N**	**%**
	English	776	94.4	6	85.7	782	94.3
	French	46	5.6	1	14.3	47	5.7
	Total	**822**	**100.0**	**7**	**100.0**	**829**	**100.0**
PGY-2+	CMG-CPG 1st	710	46.4	1365	61.2	2075	55.2
	CMG-CPG repeaters	141	9.2	164	7.4	305	8.1
	CMG-IPG 1st	1	0.1	2	0.1	3	0.1
	CMG-IPG repeaters	0	0.0	0	0.0	0	0.0
	IMG-IPG 1st	309	20.2	281	12.6	590	15.7
	IMG-IPG repeaters	186	12.2	147	6.6	333	8.9
	IMG-CPG 1st	109	7.1	179	8.0	288	7.7
	IMG-CPG repeaters	73	4.8	93	4.2	166	4.4
	Gender	**N**	**%**	**N**	**%**	**N**	**%**
	Female	769	50.3	1143	51.2	1912	50.9
	Male	760	49.7	1088	48.8	1848	49.1
	Language	**N**	**%**	**N**	**%**	**N**	**%**
	English	1237	80.9	1887	84.6	3124	83.1
	French	292	19.1	344	15.4	636	16.9
	Total	**1529**	**100.0**	**2231**	**100.0**	**3760**	**100.0**

(a) One "No Standing" candidate's results are not included in the remaining nalyses.

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 scores and domain subscore profiles are provided. Due to the small number of candidates in the PGY-1 group in the October 2019 administration, the results are suppressed in Table 6 and Figure 2.

5.1 Scale scores

Scale score summary statistics in 2019 MCCQE Part II are presented in Table 6 for the reference group (CMG-CPG 1st), total group, and PGT group.

Table 6: Summary statistics of scale scores by Candidate group

Exam session	Candidate group	PGT Group	N	Min.	Max.	Mean	Median	SD	Pass rate
May	Total	PGY-1	822	99	198	152.1	152.0	16.8	82.6
		PGY-2+	1529	50	199	146.5	150.0	24.2	69.1
	CMG-CPG 1st	PGY-1	740	100	198	153.1	153.0	16.1	85.0
		PGY-2+	710	98	199	159.4	160.0	15.5	91.4
Oct.	Total	PGY-1	7	-	-	-	-	-	-
		PGY-2+	2230	50	206	152.6	155.0	20.1	80.0
	CMG-CPG 1st	PGY-1	6	-	-	-	-	-	-
		PGY-2+	1365	92	206	159.0	159.0	15.6	91.1
Total	Total	PGY-1	829	99	198	152.1	152.0	16.9	82.5
		PGY-2+	3759	50	206	150.1	153.0	22.1	75.6
	CMG-CPG 1st	PGY-1	746	100	198	153.1	153.0	16.3	84.9
		PGY-2+	2075	92	206	159.1	160.0	15.6	91.2

Figure 1 displays the distribution of scale scores for total test takers and CMG-CPG 1st time test takers on the MCCQE Part II for May and October 2019, by PGT Group (such as PGY-1 and PGY-2+).
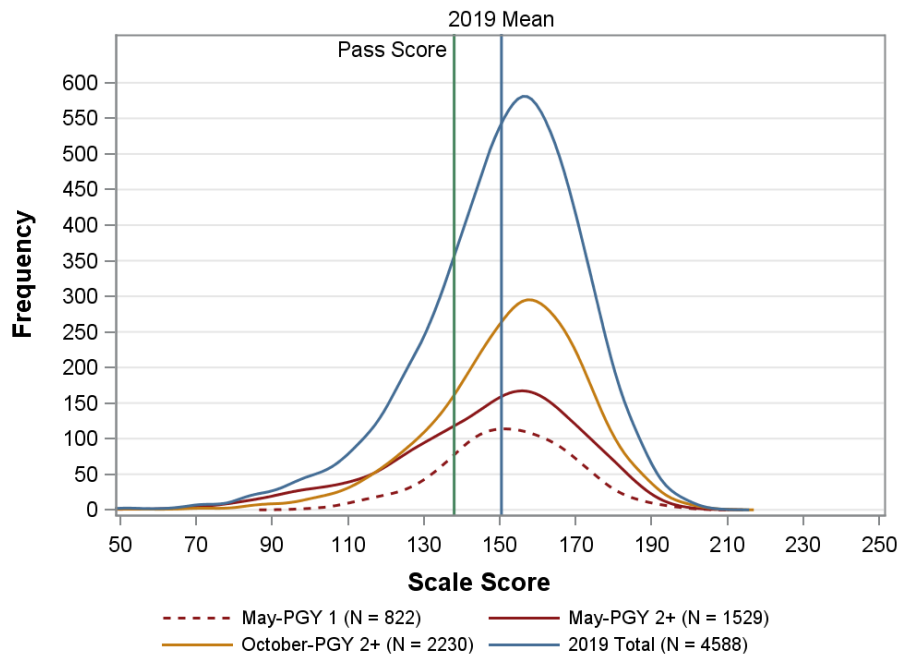

Figure 1: Scale score distribution for May and October 2019 for total candidates and by PGT group
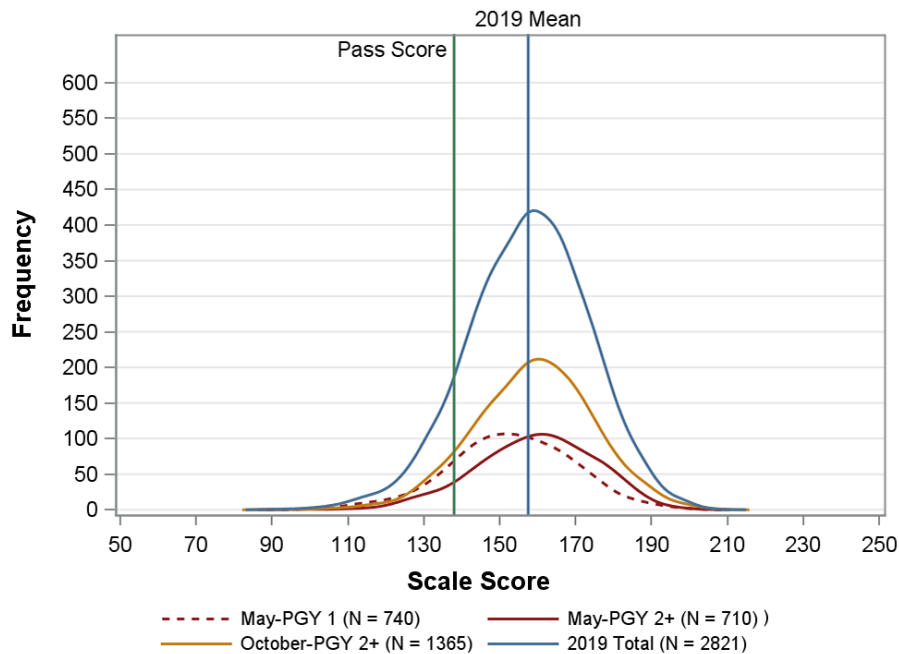
Figure 2: Scale score distribution for May and October 2019
for CMG-CPG first-time test takers, by PGT group

5.2 Estimates of reliability and classification decisions

5.2.1 Cronbach's Alpha

Cronbach's Cronbach's alpha was used to estimate score reliability for the MCCQE Part II. 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:

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

where n is the number of stations, $\sigma^2_{X_i}$ is the variance of station i scores and σ_X is the variance of total scores (Haertel, 2006, p. 74). As a general rule, a reliability estimate greater

than 0.80 on an OSCE is desirable. The reliability estimate, in conjunction with the total exam Standard Error of Measurement (SEM), provides further evidence of the reliability of the candidate's scale score.

5.2.2   Standard error of measurement

The SEM provides a value that indicates a statistical confidence range (for example, +/- 1 SEM and +/- 2 SEM represent 68 per cent and 95 per cent, respectively) within which a candidate's observed score is expected to fluctuate if the candidate was to repeat the exam over and over again. The smaller the SEM, the less measurement error in a candidate's score. 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 exam form with possible equivalent exam forms. To this end, the MCCQE Part II uses the Livingston and Lewis (1995) procedure, where decision consistency is an estimate of the agreement between classifications on potential parallel exam forms, and decision accuracy is the estimate of agreement between the observed classifications of candidates and those based on their true score (observed score ± measurement error). Ideally, both values should be high, such as 0.80 and above, supporting a reliable and valid pass/fail standing.

Table 7 shows the reliability estimates, the SEM and the decision consistency and decision accuracy estimates along with the associated false positives and false negative rates for the May and October 2019 exam forms. 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 7: Reliability estimates, SEM, decision consistency and accuracy by exam form for each administration

	May	October
Reliability estimate	0.67	0.63
SEM (score scale)	12.73	12.34
Decision consistency	0.82	0.85
False positive	0.09	0.08
False negative	0.09	0.08
Decision accuracy	0.88	0.90
False positive	0.04	0.03
False negative	0.08	0.07

It should be noted that reliability is impacted both by the amount of variability in scores amongst candidates taking a particular exam 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 exam form.

5.3 OSCE station statistics

Summary statistics for each of the OSCE stations are provided in Table 8. The percentage of missing data, average station scores or p-values, SD of station scores and Station Total Correlations (STCs) are presented. Please refer to 4.3.1 for calculation of station scores.

P-values are the average station 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 exam form.

SDs 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 (<0.30) indicates that there is a weak to negative relationship between the station score and the overall exam score. Along with the p-values, this information is 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 are flagged for content review. 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 8: Summary statistics for OSCE stations for May and October 2019

Station	May				October			
	% missing	p-value	SD	STC	% missing	p-value	SD	STC
1	0.04	0.69	0.20	0.29	0.06	0.59	0.19	0.38
2	0.11	0.67	0.18	0.33	0.06	0.58	0.16	0.23
3	0.19	0.65	0.18	0.32	0.08	0.60	0.18	0.28
4	0.07	0.72	0.19	0.31	0.07	0.62	0.21	0.35
5	0.32	0.58	0.20	0.37	0.27	0.61	0.20	0.31
6	0.11	0.72	0.20	0.38	0.04	0.73	0.19	0.35
7	0.00	0.73	0.22	0.36	0.00	0.62	0.20	0.26
8	0.08	0.64	0.16	0.27	0.07	0.75	0.19	0.20
9	0.01	0.52	0.23	0.27	0.13	0.65	0.16	0.25
10	0.00	0.58	0.19	0.38	0.07	0.77	0.15	0.29
Min.	**0.00**	**0.52**	**0.16**	**0.27**	**0.00**	**0.58**	**0.15**	**0.20**
Max.	**0.32**	**0.73**	**0.23**	**0.38**	**0.27**	**0.77**	**0.21**	**0.38**
Mean	**0.09**	**0.65**	**0.20**	**0.33**	**0.08**	**0.65**	**0.18**	**0.29**
SD	**0.10**	**0.07**	**0.02**	**0.04**	**0.07**	**0.07**	**0.02**	**0.06**

Table 8 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). 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 14-minute stations for each examiner. For the paired stations, the examiner analyses are conducted for the examiner scored component of the paired stations. The examiner analyses are based on the method outlined by Bartman, Smee and Roy (2013). For the examiner analyses, the following steps are followed:

Step One

For each examiner and station/component scored by the examiner, the average across the candidates' station scores is calculated. This average is the examiner average for that station/component. Then the average of the examiner averages is calculated along with the SD. examiners that scored fewer than 10 candidates on a station/component are excluded from these analyses as they have observed too few candidates to be compared to other examiners. Examiners are flagged as being a "Dove" if their station/component score is

higher than three times the station/component SD from the station/component average. Examiners are flagged as being a "Hawk" if their station/component score is lower than three times the station/component SD from the station/component average. For example, if the average across examiner averages was 72.5 and the SD across examiners was 6.5 and an examiner had an average of 50.7 [difference of 21.8, which is more than three SDs (6.5*3=19.5)] then he/she is flagged as a "Hawk".

Step Two

In step two, for each examiner flagged in step one, the station distribution (histogram) for the examiner is compared to the distribution of station scores from other examiners across the country. This is a visual check to evaluate whether the examiner is providing a range of scores that looks somewhat normally distributed (not providing all high or low scores). If an examiner'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 examiner flagged in step one and two, the scale-score distribution (histogram) for the cohort they scored is compared to the distribution of scale scores based on the candidates across the country. This is a check that the cohort's average scale-scores and pass rate based on all 10 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. For example, an examiner may be flagged as being a "Hawk" in steps one and two, but the candidates' scale-scores based on all 10 stations may be lower, indicating a weaker cohort. Thus, the examiner would not be flagged as a "Hawk" at step three.

Two examiners were flagged across all three steps for 2019, one May 2019 and one October 2019.

5.5 Domain subscore profiles

The purpose of the domain subscore profile is to provide feedback to candidates by highlighting their relative strengths and weaknesses. A domain subscore profile is presented in the form of two graphs for each candidate in the SIR. The graphs show the domain subscore for each of the two sets of domains (Dimensions of Care and Physician Activities) and the SEM around the domain subscore. The calculation of the domain subscores for each candidate is outlined in Section 4.3.3.

This section provides domain subscore profiles for all candidates for the May and October MCCQE Part II results. The range of domain subscores is shown graphically in Figures 3 and 4

for May 2019, and Figures 5 and 6 for October 2019. 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.
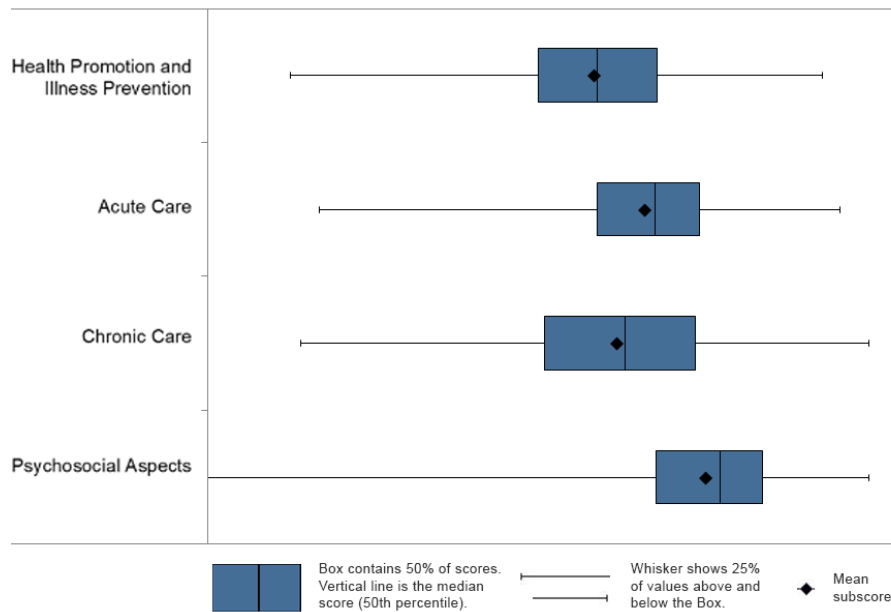

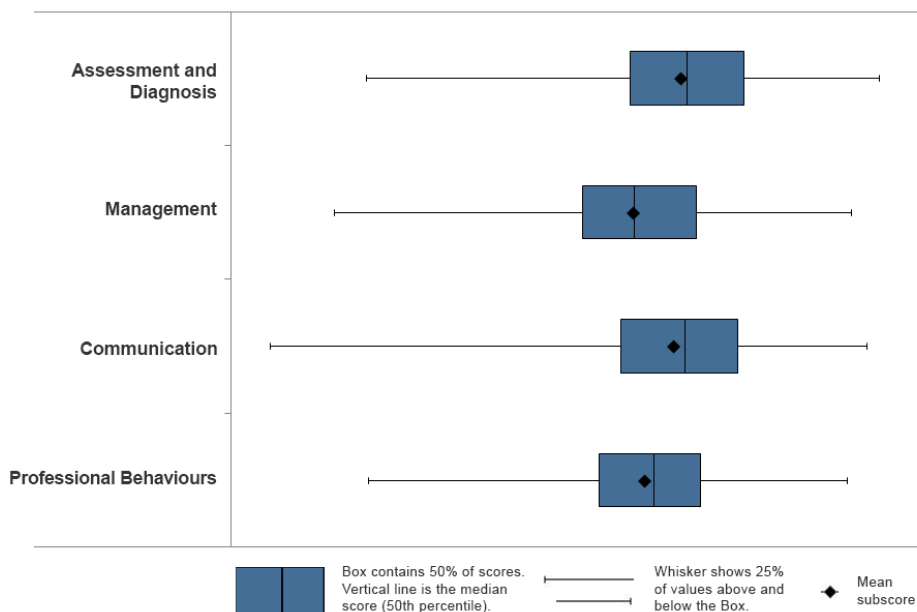Figure 3: Domain subscore for Dimensions of Care, May 2019

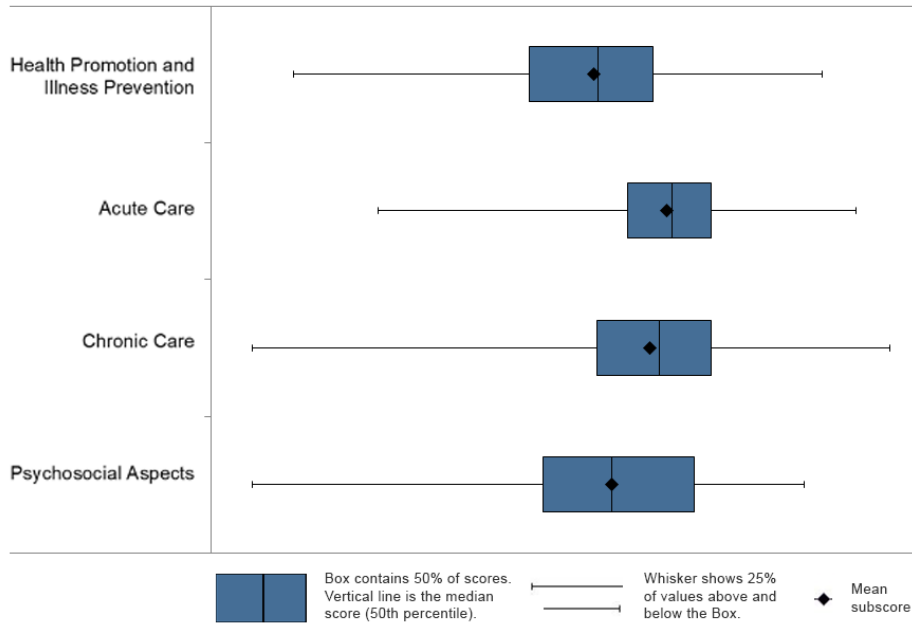Figure 4: Domain subscore profile for Physician Activities, May 2019

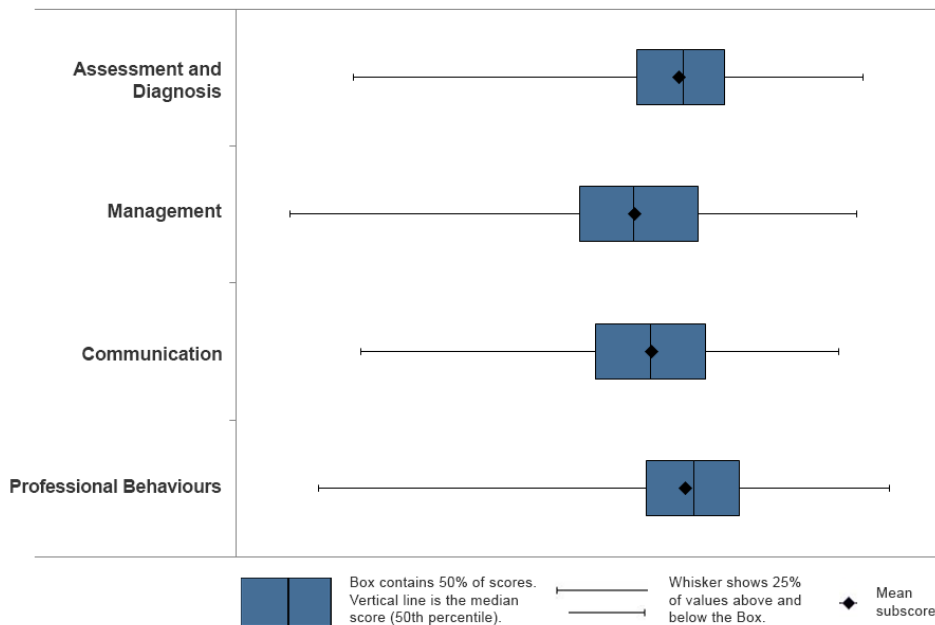Figure 5: Domain subscore for Dimensions of Care, October 2019

Figure 6: Domain subscore profile for Physician Activities, October 2019

5.6 Historical pass rates

Historical pass rates are presented in this section starting with October 2018. Table 9 shows the pass rates from October 2018[2] for four subgroups and the total group of candidates that took the MCCQE Part II by year. Table 10 show the pass rates for the May 2019 administration for the same four subgroups and total group. Table 11 show the pass rates for the October administrations 2018 to 2019 for the same four subgroups and total group. 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.

Table 9: Pass rates per calendar year by group

Year	Candidate group	N total	N pass	Pass rate
October 2018	**Total**	**2310**	**1766**	**76.5**
	CMG-CPG 1st	1495	1336	89.4
	CMG-CPG Repeat	32	22	68.8
	Other 1st	571	354	62.0
	Other Repeat	212	54	25.5
2019	**Total**	**4588**	**3524**	**76.8**
	CMG-CPG 1st	2821	2526	89.5
	CMG-CPG Repeat	305	259	84.9
	Other 1st	962	529	55.0
	Other Repeat	500	210	42.0

Table 10: Pass rates by May administrations by group

Year	Candidate group	N total	N pass	Pass rate
2019	**Total**	**2351**	**1734**	**73.8**
	CMG-CPG 1st	1450	1278	88.1
	CMG-CPG Repeat	141	116	82.3
	Other 1st	500	252	50.4
	Other Repeat	260	88	33.8

Table 11: Pass rates by October administrations by group

Year	Candidate group	N total	N pass	Pass rate
2018	**Total**	**2310**	**1766**	**76.5**
	CMG-CPG 1st	1495	1336	89.4
	CMG-CPG Repeat	32	22	68.8
	Other 1st	571	354	62.0
	Other Repeat	212	54	25.5
2019	**Total**	**2237**	**1790**	**80.0**
	CMG-CPG 1st	1371	1248	91.0
	CMG-CPG Repeat	164	143	87.2
	Other 1st	462	277	60.0
	Other Repeat	240	122	50.8

---

[2] The pass rates for May 2018 are not included in this report as a new exam blueprint was implemented with the October 2018 administration.

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.

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*

Medical Council of Canada. (2019). *Technical report on the standard setting exercise for the Medical Council of Canada Qualifying Examination Part II*. MCC: Ottawa, Canada. Retrieved from: *mcc.ca/media/MCCQE-Part-II-Standard-setting-report-2018.pdf*

APPENDIX A:
QUALITY CONTROL – MCCQE PART II RESULTS

Pre-examination

1.	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 language and perform quality assurance (QA) checks of raw data in CSV files	
	QEII Production Coordinator	(Initials & sign-off date)
2.	Create answer keys, compare against exam sheets, case, scoring rules and adjust as necessary;	
	QEII Production Coordinator	(Initials & sign-off date)
3.	Import answer key into SQL database and download answer key to SAS	
	Analyst (SQL)	(Initials & sign-off date)
	Analyst (SAS)	(Initials & sign-off date)
4.	Update the SAS code as necessary (new question type, etc.)	
	Analyst (SAS)	(Initials & sign-off date)
5.	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	(Initials & sign-off date)
	Analyst (SQL)	(Initials & sign-off date)
	Analyst (SAS)	(Initials & sign-off date)
	Senior Research Psychometrician (Overview of analyses)	(Initials & sign-off date)
6.	Test candidate bar code labels	
	QEII Production Coordinator	(Initials & sign-off date)
7.	Test examiner bar code labels	
	Administration Coordinator	(Initials & sign-off date)
8.	Prepare a file with demographic information about candidates including creating candidate groups (candidate info)	
	Analyst (SQL)	(Initials & sign-off date)
	Analyst (SAS)	(Initials & sign-off date)
	Senior Research Psychometrician (Overview of analyses)	(Initials & sign-off date)

Post-examination (Pre-CEC)

These tasks _MUST_ be completed before the CEC meets.

9.	On-site scanning, receiving and uploading data files from sites and load files into TeleForm®	
	QEII Production Coordinator	(Initials & sign-off date)
10.	Verify	
	CSA team	(Initials & sign-off date)
11.	Import CSVs into scoring application and correct until error reports are clean	
	QEII Production Coordinator	(Initials & sign-off date)
	Analyst (SQL)	(Initials & sign-off date)
12.	Run SAS scoring in parallel with scoring in SQL; verify matching results in SAS and scoring application	
	Analyst (SQL)	(Initials & sign-off date)
	Analyst (SAS)	(Initials & sign-off date)
13.	Run pull lists and QA against SAS	
	QEII Production Coordinator	(Initials & sign-off date)
	Production Coordinator	(Initials & sign-off date)
	Analyst (SAS)	(Initials & sign-off date)
14.	Pull sheets, hand score, enter into hand scoring application and update CSV files	
	QEII Production Coordinator	(Initials & sign-off date)
15.	Run preliminary statistical analyses	
	Analyst (SAS)	(Initials & sign-off date)
16.	Review statistical information	
	Senior Research Psychometrician	(Initials & sign-off date)
17.	Present preliminary statistics to the Manager, MCCQE Part II and Associate Director, Evaluation Bureau	
	Senior Research Psychometrician	(Initials & sign-off date)
18.	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	(Initials & sign-off date)
	Note unusual pattern: * Add any supporting documentation to the folder. * Additional notes or comments can be added to the last page of this document.	

APPENDIX B:
PATIENT INTERACTION RATING SCALE ITEMS

1.0 Attentiveness to Ethical Issues (as relevant to this case) [**One** bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Exhibits professional behaviours including compassion, respect and maintenance of confidentiality o Addresses ethical issues (e.g., disclosure and consent) with clarity and respect o Demonstrates awareness of and adherence to regulatory and legal requirements	O	O	O	O

2.0 Interviewing Skills [**One** bubble only] [**One** bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Gathers relevant biomedical and psychosocial information o Establishes a timeline or sequence of events o Elicits pertinent positives and negatives as they relate to the differential diagnosis o Uses open and closed-ended questions o Avoids jargon and leading questions o Attentive to verbal and non-verbal cues o Clarifies and summarizes what the patient has said	O	O	O	O

3.0 Patient-Centred Approach to History-Taking / Physical Examination [**One** bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Explores patient's experience with health problem (e.g., feelings, ideas, impact and expectations) o Attends to patient's verbal and/or physical responses o Attempts to understand patient's psychosocial context	O	O	O	O

4.0 Patient-Centred Approach to Treatment Planning [**One** bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Engages patient in healthcare planning (e.g., encourages questions, discussion and feedback) o Integrates treatment approach with patient's psychosocial context and priorities o Encourages adherence to treatment by finding common ground	O	O	O	O

5.0 Patient Education [**One** bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Provides clear, concise and accurate information o Provides a summary of expected disease course, progression or resolution o Verifies that the information is understood o Avoids jargon o Facilitates informed decision-making by exploring risks and benefits of each option o Discusses timeline for follow-up	O	O	O	O

6.0 Organization of Encounter [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Pursues a purposeful encounter with a logical flow o Explores the most pertinent data; does not lose time on less relevant information o Intervenes with the patient as appropriate	O	O	O	O
7.0 Interpersonal Behaviour [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Demonstrates respectful management of the interaction (e.g., non-judgmental, culturally sensitive, avoids interrupting) o Listens appropriately to facilitate conversation o Uses appropriate body language; remains composed o Avoids offensive or aggressive behaviour	O	O	O	O
8.0 Approach to Physical Examination [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Selects appropriate manoeuvers for the patient problem o Performs examination in a logical order o Demonstrates technical skill o Executes the examination respectfully; elicits and interprets significant findings o Attentive/Shows attentiveness to the patient's physical comfort and dignity	O	O	O	O
9.0 Patient Summary Report: Organization [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Report is well organized, coherent, and logical o Reflects a complete information transfer	O	O	O	O
10.0 Patient Summary Report: Key Issues [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Report reflects comprehensive understanding of key issues o Report includes pertinent negative/normal findings, in addition to the positive/abnormal issues o Communication is concise with no irrelevant information	O	O	O	O
11.0 Patient Summary Report: Assessment and Management Plan [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Reports a well-reasoned plan that considers impact on patient o Provides rationale for treatment	O	O	O	O
12.0 Collaboration (non-patient / colleague) [One bubble only]	Inadequate performance	Marginal performance	Adequate performance	Superior performance
o Engages respectfully (verbal and non-verbal) o Discusses pertinent information o Shares responsibility for safe and effective patient care o Manages differences and resolves conflicts to enhance collaboration o Manages differences and resolves conflicts to enhance collaboration	O	O	O	O

APPENDIX C: EXAMPLE SOR AND SFR

Medical Council of Canada
Qualifying Examination Part II
Statement of Results

MEDICAL COUNCIL   LE CONSEIL MÉDICAL
OF CANADA   DU CANADA

Candidate name: Xxxxxxxxxxxx, Xxxxxxxx

Candidate code: 0000000000

Examination session: May 2019   **Your final result:**   Pass

Pass score: 138   **Your total score:**   240

July 6, 2019

We are writing to inform you of your final result on the Medical Council of Canada Qualifying Examination Part II.

Your total score is reported as a scaled score ranging from 50 to 250 with a mean of 150 and a standard deviation of 20. The mean and standard deviation were set using the results from the May 2019 session.

Your final result is based on your total score relative to the pass score.

For more information, please visit the exam's Scoring web page on our website, mcc.ca.

Supplemental information on your examination performance is reported to you in a separate document within your physiciansapply.ca account.

mcc.ca
physiciansapply.ca
inscriptionmed.ca

Medical Council of Canada
Qualifying Examination Part II
Supplemental Information Report

Candidate name:	Xxxxxxxxxxxx, Xxxxxxxx
Candidate code:	0000000000
Examination session:	May 2019

Your final result:	Pass
Your total score:	240

This report provides you with supplemental information on your performance on the Medical Council of Canada Qualifying Examination (MCCQE) Part II.

The MCCQE Part II assesses core abilities to apply medical knowledge, demonstrate clinical skills, develop investigational and therapeutic clinical plans, as well as demonstrate professional behaviours and attitudes at a level expected of a physician in independent practice in Canada.

The exam assesses your performance across two broad categories with each exam question classified on both categories:
- Dimensions of care, covering the spectrum of medical care
- Physician activities, reflecting a physician's scope of practice

Each category has four domains:

Dimensions of Care	Physician Activities
Health Promotion and Illness Prevention	Assessment and Diagnosis
Acute Care	Management
Chronic Care	Communication
Psychosocial Aspects	Professional Behaviours

Figure 1 displays your performance in each domain under Dimensions of Care. Figure 2 displays your performance in each domain under Physician Activities.

In both figures, we provide your subscores along with the mean subscore of first-time takers who passed the exam in the same session.

Each domain is assigned a weighting on the exam. We present the content weights, expressed as percentages, in the grids shown on page 3.

We also provide the standard error of measurement (SEM) for each of your subscores. It represents the expected variation in your subscore if you were to take this exam again with a different set of stations covering the same domains.

Small differences in subscores or overlap between SEMs indicate that performance in those domains was somewhat similar. Overlap between the SEM and the mean score of first-time takers who passed signifies that performance is similar to the mean.

Subscores are based on less data than the total score and have less precision. Your total score and subscores cannot be compared as they are calculated differently. The pass score cannot be applied to Figure 1 or 2.

For more information, please visit the exam's Scoring web page on our website mcc.ca.

Figure 1: Dimensions of Care

YOUR PERFORMANCE

- Health Promotion and Illness Prevention
- Acute
- Chronic
- Psychosocial Aspects

Low performance — High performance

Mean subscore of first-time takers who passed

SEM: Standard error of measurement
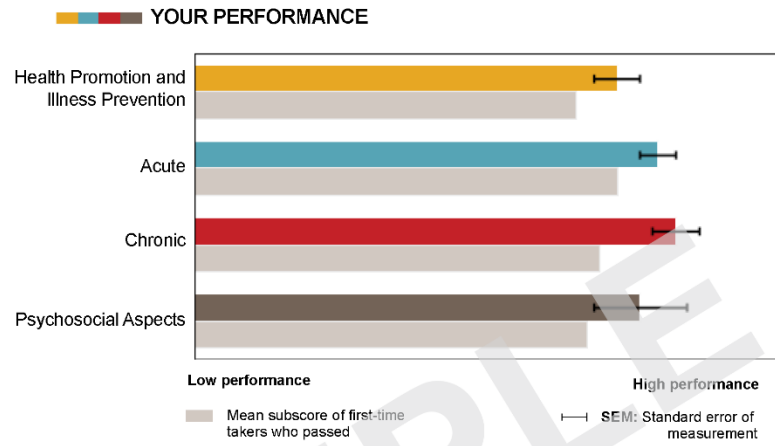
Figure 2: Physician Activities

YOUR PERFORMANCE

- Assessment and Diagnosis
- Management
- Communication
- Professional Behaviours

Low performance — High performance

Mean subscore of first-time takers who passed

SEM: Standard error of measurement

Report: July 6, 2019
Candidate code: 0000000000

2/3

Dimensions of Care

Reflects the focus of care for the patient, family, community and/or population:

- **Health Promotion and Illness Prevention:** The process of enabling people to increase control over their health and its determinants, and thereby improve their health. Illness prevention covers measures not only to prevent the occurrence of illness, such as risk factor reduction, but also to arrest its progress and reduce its consequences once established. This includes, but is not limited to screening, periodic health exam, health maintenance, patient education and advocacy, and community and population health.

- **Acute:** Brief episode of illness within the time span defined by initial presentation through to transition of care. This dimension includes but is not limited to urgent, emergent, and lifethreatening conditions, new conditions, and exacerbation of underlying conditions.

- **Chronic:** Illness of long duration that includes but is not limited to illnesses with slow progression.

- **Psychosocial Aspects:** Presentations rooted in the social and psychological determinants of health and how these can impact on wellbeing or illness. The determinants include but are not limited to life challenges, income, culture, and the impact of the patient's social and physical environment.

Physician Activities

Reflects the scope of practice and behaviours of a physician practicing in Canada:

- **Assessment/Diagnosis:** Exploration of illness and disease using clinical judgment to gather, interpret and synthesize relevant information that includes but is not limited to history taking, physical examination and investigation.

- **Management:** Process that includes but is not limited to generating, planning, organizing safe and effective care in collaboration with patients, families, communities, populations, and other professionals (e.g., finding common ground, agreeing on problems and goals of care, time and resource management, roles to arrive at mutual decisions for treatment, working in teams).

- **Communication:** Interactions with patients, families, caregivers, other professionals, communities and populations. Elements include but are not limited to relationship development, intra-professional and inter-professional collaborative care, education, verbal communication (e.g., using the patient-centered interview and active listening), non-verbal and written communication, obtaining informed consent, and disclosure of patient safety incidents.

- **Professional Behaviours:** Attitudes, knowledge, and skills relating to clinical and/or medical administrative competence, communication, ethics, as well as societal and legal duties. The wise application of these behaviours demonstrates a commitment to excellence, respect, integrity, empathy, accountability and altruism within the Canadian health-care system. Professional behaviours also include but are not limited to self-awareness, reflection, life-long learning, leadership, scholarly habits and physician health for sustainable practice.

3/3