FM-MAP: A Novel In-Training Examination Predicts Success on Family Medicine Certification Examination

Karl Iglar, MD; Fok-Han Leung, MD; Rahim Moineddin, PhD; Jodi Herold, PhD

BACKGROUND AND OBJECTIVES: The objective of our study was to assess the correlation between a locally developed In-Training Examination (ITE) and the certification examination in family medicine in Canada.

METHODS: The ITE was taken twice yearly, which corresponded for most residents to the fifth, ninth, 17th, and 21st month of training. The results for the ITE were correlated to the CFPC certification examination taken in the 23rd month of residency.

RESULTS: The scores on each of the four iterations of the ITE correlated moderately well with performance relating to problem solving skills and knowledge on the certification examination. The ITE showed a trend to an increased correlation with duration in the training program with a Spearman correlation coefficient increasing from 0.45 on the first test to 0.54 on the fourth test. The correlation of the ITE with performance on the component assessing the doctor-patient relationship on the certification examination was poor (r=0.26 on the last test).

CONCLUSION: Our in-training examination is a useful predictor of performance in problem solving and knowledge domains of the family medicine expert role on the certification examination.

(Fam Med. 2017;49(5):369-73.)

Current methods used to assess residents for readiness to take the certification examination fail to identify a significant proportion of unprepared residents. Program directors are only able to accurately identify 25% of the top and bottom performers on the examination. In addition, faculty have difficulty accurately assessing the knowledge base of trainees they supervise, and residents are unable to predict their own performance on knowledge-based tests.

In-Training Examinations (ITE), which follow a similar testing format (multiple choice questions) to certification examinations, have been shown to be predictive of success on certification examinations in a number of disciplines, including anesthesiology, internal medicine, and family medicine. Currently, in-training examinations (ITE) in family medicine are not the norm in Canada. There is also little information regarding the predictive value with respect to certification examination results of a locally developed ITE or of an ITE that does not utilize the same format.

The Department of Family and Community Medicine (DFCM) at the University of Toronto has one of the largest training programs in family medicine in Canada, with over 350 residents enrolled in the 2-year program. The residents are assigned to one of 14 geographically distributed sites as well as a Rural Residency Program. Based on the learning outcomes of the Competency-Based Curriculum in Family Medicine at the University of Toronto, the program has implemented an iterative ITE known as the Family Medicine-Mandatory Assessment of Progress (FM-MAP), using a multiple-choice multiple-answer format to assess knowledge acquisition and application components of the family medicine expert role. Each question on the FM-MAP is mapped to learning outcomes in the curriculum that are used in DFCM to define practice readiness (see Table 1). Although addressing these learning outcomes in training will prepare residents for the certification examination, they do not serve as the blueprint for the content of this examination.

The objective of this study was to determine if the FM-MAP, which uses a different question format than the certification examination (multiple choice versus short answer questions), could predict performance on the high-stakes certification examination. The multiple choice question (MCQ) format was chosen due
to ease of marking and therefore the ability to provide feedback in a timely way to residents.

Methods
Participants
There were 132 postgraduate year (PGY)-2 trainees deemed ready to take the certification examination in May 2012. A total of 124 of these residents took the FM-MAP in March 2012 and were the subjects for this study.

Instruments
The FM-MAP is used to assess knowledge application. The residents take the FM-MAP twice yearly, which corresponds for most residents to the fifth, ninth, 17th, and 21st month of training. It is administered during weekly protected academic time and is comprised of over 120 questions. The questions in the FM-MAP are different for each iteration of the test. The scores are generated as percentiles, which are then norm-referenced and reported as decile ranks. Evidence of construct validity of the FM-MAP across a number of parameters (PGY level, source of medical degree, clinical exposure, and comparison of top versus bottom performers based on program director assessment) has been described previously.11

The College of Family Physicians of Canada (CFPC) administers the certification examination to family medicine residents who have successfully completed at least 18 months of training.12 The examination has two components: the written short answer management problems (SAMPs) and simulated office oral (SOO) stations.13 There are 40–45 SAMPs, and they are intended to assess problem-solving skills and knowledge in the context of a clinical situation.14 The five SOO stations are intended to assess the doctor-patient interaction and communication.15,16 Each resident obtains an aggregate score for the SAMPs and an aggregate score for the SOOs. In order to obtain certification, a resident must obtain a passing aggregate score on both components. The passing score is determined using the aggregate scores of graduates of Canadian family medicine residency programs who are taking the examination for the first time.17 A Z-score of less than -2 on either component indicates a failure.

Analysis
The Spearman correlation was used to assess the correlation between the FM-MAP and the certification examination SAMP and SOO components for the four iterations. The linearity of the relationship between the FM-MAP and SAMP and SOO components was assessed using a scatter plot. Sensitivity, specificity, positive predictive values, and negative predictive values were calculated for the FM-MAP.

The study was approved by the Research Ethics Board of the Faculty of Medicine at the University of Toronto.

Results
There were 124 residents who took the FM-MAP in March 2012 and the certification examination in May 2012 and included 98 residents who took all four iterations of the FM-MAP. Of these 124 residents, 71 (57.3%) were female and 53 (42.7%) were male, with 103 (83.1%) being Canadian medical graduates and 21 (16.9%) being international medical graduates.

The Spearman correlation coefficients for the four iterations of the FM-MAP and the certification examination SAMP and SOO components were calculated and are presented in Table 1. The Pearson correlation coefficients were calculated for the FM-MAP and the certification examination SAMP and SOO components for the four iterations and are presented in Table 2.

Table 1: A Sample FM-MAP Case Showing Learning Outcomes Assessed

<table>
<thead>
<tr>
<th>Case</th>
<th>FAM10. Demonstrate an effective approach to the diagnosis and management of common chronic diseases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are a family doctor working at a busy urban family medicine clinic. John is a 30-year-old male who you have known for 2 years. He is a general laborer with a Grade 10 education. He has no specific past medical history and is on no prescribed medications. Since you have known him he has always complained about his back pain of several years duration. In the past you have been diligent in taking a history and doing a thorough physical examination and have confidently diagnosed him as having “mechanical chronic low back pain.” He has told you in the past that he has been taking over-the-counter acetaminophen, sometimes with codeine 8 mg, to help relieve the pain. Today he breaks down in tears in the office saying that his pain has been quite severe over the past 3 months, that he has increased his acetaminophen with codeine 8 mg intake to 10–15 tablets per day and that he “just can’t take the pain anymore.” The pain is still just in his low back. There is no pain in the buttocks, hips, or legs. There is no numbness or weakness. John tells you that his mood has been depressed for the past 3 weeks.</td>
<td>FAM11. Demonstrate an effective approach to the presentation of illnesses with a strong psychological component.</td>
</tr>
</tbody>
</table>
examination SAMP and SOO components for our cohort are shown in Table 2. All correlations were statistically significant with \( P < 0.05 \). The examination of the scatter plot (shown in Figure 1) confirmed a linear relationship between the FM-MAP and the certification examination.

There were 11 residents who failed the certification examination in May 2012 and took the FM-MAP in March 2012. Five residents failed the written SAMP component, five failed the SOO component, and one failed both.

In terms of the FM-MAP, 46 residents were in the 4th or lower decile in terms of performance, and 78 were in the 5th decile or above. Of the residents who failed the SAMP component, five were in the 4th decile or below on the FM-MAP, and one was in the 5th decile or above. Using the arbitrary cut-off of the 4th decile, the sensitivity, specificity, positive predictive value, and the negative predictive value were calculated and are presented in Table 3.

**Discussion**

The study was a retrospective analysis of performance of residents on the certification examination compared to their performance on in-training examinations. There was a moderate statistically significant correlation between each of the four iterations of the FM-MAP and the SAMP component of the certification examination. The strength of the correlation increased with each iteration and hence time in the program ranging from 0.45 on the first FM-MAP to 0.54 on the fourth and final FM-MAP. Not surprisingly, the correlation between the last FM-MAP and the SOO component of the certification examination was poor at 0.27, and it was even lower for prior iterations of the test. This poor correlation is likely related to the fact that this component of the examination has been utilized to assess the doctor-patient interaction and not knowledge application, which both the SAMPs and the FM-MAP address. This demonstrates that when similar elements such as knowledge acquisition and application are tested by in-training and certification examinations, there is a correlation between the tests. A resident whose score was in the 4th decile or below

---

### Table 2: Correlation Coefficients* Between the FM-MAP and the SAMP and SOO Components of the Certification Examination

<table>
<thead>
<tr>
<th></th>
<th>November 2010</th>
<th>April 2011</th>
<th>November 2011</th>
<th>March 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMP</td>
<td>0.45</td>
<td>0.46</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>SOO</td>
<td>0.15</td>
<td>0.14</td>
<td>0.19</td>
<td>0.27</td>
</tr>
</tbody>
</table>

* The correlation was determined using the Spearman Rank Correlation Coefficient, and all correlations were statistically greater than 0.
Table 3: Predictive Ability of the Last FM-MAP Using a 4th Decile Cut-Off

<table>
<thead>
<tr>
<th>SAMP (Certification)/FM-MAP</th>
<th>Z score &lt; -2 (Fail)</th>
<th>Z score ≥ -2 (Pass)</th>
<th>Total</th>
<th>Positive predictive value</th>
<th>5/46 = 11%</th>
<th>Negative predictive value</th>
<th>77/78 = 99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decile ≤ 4</td>
<td>5</td>
<td>41</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decile ≥ 5</td>
<td>1</td>
<td>77</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>118</td>
<td>124</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>5/6 = 83%</td>
<td>Specificity</td>
<td>77/118 = 65%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

on the final FM-MAP was 8.4 times more likely to fail the short answer management problem (SAMP) component of the CFPC certification examination. Therefore, achieving a score in the 5th decile or above on the last FM-MAP resulted in a 99% likelihood of passing the SAMP component of the certification examination. The results of the last iteration of the FM-MAP may identify a group of residents who are at risk of being unsuccessful on the SAMP component of the certification examination and may allow for an opportunity to intervene to better prepare residents for the examination.

A recent study reported a correlation of 0.7 between the American Board of Family Medicine (ABFM) in-training examination and the ABFM certification examination. The in-training examination in this study is “built to the same specifications” as the certification examination. It must be noted that the question format of the FM-MAP and the certification examination is different. The FM-MAP follows a multiple-choice multiple-answer format whereas the CFPC certification examination SAMP component follows a short answer format. This likely resulted in a lower correlation coefficient in our study, but the correlation is consistent with the literature comparing written assessments utilizing different question formats.

An additional factor may be that the timing of the FM-MAP and the certification examination is asynchronous. Residents receive results of the FM-MAP more than 2 months prior to the certification examination. This may lead to a weaker correlation coefficient due to the fact that it may motivate residents performing poorly on our test to alter their study habits. This is an area that will require further study.

Our study builds on the current literature describing correlation between in-training and certification examinations. The difference is that our study shows the correlation for examinations that do not have the same question formats. This allows programs to either develop their own tests or use existing in-training examinations even though the question format may be different when compared to the discipline’s certification examination.

Limitations
The study represents a correlation between performance on the FM-MAP and the certification examination by one cohort of 124 residents. The four iterations of the FM-MAP are not equated across administrations. The FM-MAP scores used as a comparison to the certification examination components represent decile rankings. Similarly the certification examination scores for the SAMP and SOO components are reported as Z-scores. The use of norm-referencing for both the FM-MAP and the certification examination may lead to different results if data from future cohorts are compared.

Conclusions
Identifying residents at risk of being unsuccessful on certification examinations remains a challenge for program directors and faculty. Residents themselves have been shown to be poor assessors of their own knowledge. ITE help identifies those residents who may be unprepared to take the certification examination as they correlate well with certification examinations when they both assess knowledge acquisition and application. This occurs even when the question formats of the respective examinations are different. Programs can create or use ITE specific to their context and be confident in their utility irrespective of the question format of their certification examinations.

CORRESPONDING AUTHOR: Address correspondence to Dr Iglar, Department of Family and Community Medicine, University of Toronto, St. Michael’s Hospital, 500 University Avenue, 3rd Floor, Toronto, Ontario, Canada, M5G 1V7. 416-978-6467. Fax: 416-978-8179. kari.iglar@utoronto.ca.

References


