

ORIGINAL RESEARCH

Physician Opinions about American Board of Family Medicine Self-Assessment Modules (2006–2016)

Mingliang Dai, PhD, Michael Hagen, MD, Aimee R. Eden, PhD, MPH, and Lars E. Peterson, MD, PhD

Introduction: Maintenance of Certification (MOC) was implemented to help physicians remain current with evolving medical standards, but has been criticized for being irrelevant to practice. We assessed family physicians' (FPs') opinions about the content of American Board of Family Medicine (ABFM) self-assessment modules (SAMs).

Methods: We used ABFM administrative data from feedback surveys completed after each of the 16 SAMs from 2006 to 2016. FPs rated agreement with 2 statements—1) "Content is appropriate for my practice," and 2) "Content was presented at an appropriate level"—on a scale of 1 (strongly disagree) to 6 (strongly agree). We calculated mean ratings of each statement by year and stratified by Knowledge Assessment (KA) and Clinical Simulation (CS) portions of the SAM. We plotted mean ratings by FPs' age at their first SAM completion and the total number of SAMs completed.

Results: SAMs were completed ($n = 633,198$) from 2006 to 2016 with 448,408 (71%) feedback surveys completed. The annual mean ratings of both statements varied little (less than 0.5) and were above 4.5 for all SAMs. CS ratings were consistently lower than KA ratings. FPs of all ages at first SAM provided similar ratings and agreement with content appropriateness increased with repeated exposure to SAMs.

Conclusion: Over 11 years, the content of ABFM SAMs was regarded by FPs as appropriate for practice and presented at an appropriate level. Continued monitoring of feedback is necessary to keep the content of MOC programs relevant for physicians' practice. (J Am Board Fam Med 2019;32:79–88.)

Keywords: Certification, Family Physicians, Self-Assessment, Surveys and Questionnaires

The American Board of Medical Specialties (ABMS) approved a Maintenance of Certification (MOC) program in 2000 to provide a platform to continuously assess physicians' performance and currency with evolving medical evidence. ABMS set standards for MOC that allowed each member board to design a program to meet the needs of their Dip-

lomates within a 4-part framework: professional standing (Part I), lifelong learning and self assessment (Part II), medical knowledge and skills (Part III), and improvement in practice (Part IV).¹ However, these standards have sometimes triggered tensions between certifying boards and their Diplomates.² Despite growing evidence linking MOC participation with improvement in patient care,^{3–8} opinions about MOC have been controversial and its value challenged.^{9–11} Physician experiences with MOC have demonstrated both satisfaction,^{12–15} and dissatisfaction.^{16,17}

The ABMS MOC Part II activities are intended to be relevant, easy-to-use, cost-effective, and meaningful for physicians.¹ However, there is little evidence demonstrating whether these goals have been met. The American Board of Family Medicine (ABFM), as an ABMS member board, is able to

This article was externally peer reviewed.

Submitted 22 December 2017; revised 19 June 2018; accepted 20 June 2018.

From the American Board of Family Medicine, Lexington, KY (MD, MH, ARE, LEP); Department of Family Community Medicine, University of Kentucky, Lexington (MH, LEP).

Funding: none.

Conflict of interest: The authors are employees of the American Board of Family Medicine.

Corresponding author: Mingliang Dai, PhD, 1648 McGrathiana Parkway, Suite 550, Lexington KY, 40509 (E-mail: mdai@theabfm.org).

offer insights regarding physician opinions about its MOC for Family Physicians (MC-FP) Part II activities, as a result of its long-standing business requirement of collecting feedback from Diplomates who complete a self-assessment module (SAM), which comprised a knowledge assessment (KA) and a clinical simulation (CS). Soon after their rollout in 2004, an early analysis of the ABFM feedback surveys indicated that most family physicians (FPs) had a favorable experience with SAMs,¹³ and a more recent analysis of over 320,000 feedback surveys of FPs who completed a SAM between 2004 and 2013 found that a large majority of participants rated their experience with SAMs as highly positive.¹⁸

As the potential scope of practice of FPs is very broad, it is important to determine whether the SAMs, which are focused on a specific clinical area, are providing knowledge that is appropriate for practice. Scope of practice often changes as FPs age and an internal ABFM analysis of 6 months of data found that older FPs rated KA more positively than CS compared with younger FPs.¹⁹ As a result, we hypothesized that older Diplomates may have found the CS difficult and been less satisfied than younger Diplomates. This study sought to evaluate FPs' opinions about ABFM SAMs and also to determine whether opinions differed between younger and older FPs.

Methods

Data

This study analyzed data from ABFM SAM feedback surveys completed between January 2006 and June 2016 before ABFM decoupled SAMs into independent Knowledge Self-Assessment (KSA) and Clinical Self-Assessment (CSA) activities.²⁰ Until 2009, physicians were required to complete the surveys to receive Continuing Medical Education (CME) credits. After 2009, survey completion was optional.

During the period of the study, participants in MC-FP needed to complete 2 SAMs in a 3-year period to satisfy the lifelong learning requirement. Each SAM included 2 components: a 60-question Knowledge Assessment (KA) that focused on medical knowledge, and a Clinical Simulation (CS) to simulate a patient clinical encounter. To keep the SAMs current, ABFM adopted multiple strategies to update the content regularly.²¹ Diplomates had to answer 80% of the KA questions correctly be-

fore proceeding to the CS. After completing both SAM components, FPs were asked to evaluate the activities. For each KA and each CS, FPs were asked to rate how much they agreed with the following statements: 1) "Content is appropriate for my practice" and 2) "Content was presented at an appropriate level." The number of SAM topics from which Diplomates could choose increased from 2 (Diabetes and Hypertension), at launch in 2004 to a selection of 16 topics by 2016. The 2 questions about content appropriateness were first asked in 2006 and remained unchanged in all 16 SAMs through 2016.

Measures

Statements of content appropriateness were rated on a numeric scale of 1 to 6 (from 1, strongly disagree to 6, strongly agree) for most SAMs. For the 4 SAMs (Pain Management, Well Child Care, Health Behavior, Maternity Care) that employed a rating scale of 1 to 4, we averaged the ratings using the raw scores and then weighted the mean ratings by a factor of 1.5 to align them with the 6-point scale. Mean ratings of the statements were used to measure FPs' opinions about content appropriateness of individual SAM KA/CS. Overall opinions about KA/CS were measured by aggregating mean ratings across all SAMs to KA/CS level. In addition, age at first SAM completion was used to establish cohorts of FPs aged under 40 years, 40 to 49 years, 50 to 59 years, and 60 years and older. Total number of SAMs completed during the study period was counted for each physician.

Analysis

We calculated the percentage of feedback surveys completed by SAM topic and by year. To test survey response bias, we examined whether respondents and nonrespondents differed in age, gender, medical degree (MD vs DO), and International Medical Graduate (IMG) status using data from 2015, the most recent year with a full year's data. We compared the differences within each SAM rather than across SAMs within a calendar year because it became difficult to differentiate respondents from nonrespondents when a physician could complete multiple SAMs in a year but responded to any or none of the surveys. We examined mean ratings of content appropriateness for each SAM KA/CS topic over the years of its offering. Lastly, we investigated differences in overall opinions

about content appropriateness of KA/CS by plotting the aggregated mean ratings against the total number of SAMs completed and by physician age cohort. The population of all SAM feedback surveys completed between 2006 and 2016 were available for analysis, mitigating the need for inferential statistical testing. Analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC) and Microsoft Excel. This study was approved by the American Academy of Family Physicians Institutional Review Board.

Results

Over the 11-year period, 448,408 feedback surveys (from 633,198 SAMs) were completed for a response rate of 71%. The number of SAMs completed each year by topic can be found in the Appendices. The response rates before 2009 were above 85% for the 10 SAMs that were available and, expectedly, declined after the requirement for CME credits was removed (Table 1). After reaching the lowest point in 2012 when only 56% of SAM feedback surveys were completed, the response rates started to increase. Of all the SAMs, the mean response rate was highest (77%) with the Asthma SAM and lowest with the Preventive Care SAM (59%) during the study period.

Compared with nonrespondents, feedback survey respondents in 2015 were not significantly different in gender, medical degree, or IMG status for 10, 13, and 9 of the 16 SAMs respectively (see complete results in the Appendices). In contrast, the mean age was higher among respondents for all the 16 SAMs except for Diabetes (45.9 vs 45.4, $P = .09$) and Preventive Care SAMs (44.5 vs 45.1, $P = .06$). The largest mean age difference was 3.1 years (Mental Health, 43.2 vs 46.3, $P < .0001$).

Within individual KA modules, changes in ratings over the years were minimal (eg, the largest mean difference was 0.3 between any 2 years) and all the KAs were rated higher than 4.5 (of 6) (Tables 2 and 3). The mean ratings for “Content is appropriate for my practice” and “Content was presented at an appropriate level” were nearly identical. Across modules, some KAs (eg, Coronary Artery Disease and Cerebrovascular Disease) were rated slightly lower than others (eg, Diabetes or Hypertension). Nevertheless, the largest mean rating difference in any given year was less than 0.5. The overall pattern of the ratings within and across modules held in CSs (Tables 4 and 5). Compared with KAs of the same topic area, most CSs were rated lower but only by a magnitude of 0.1 to 0.2, with the exception of the Maternity Care and

Table 1. Response Rates to American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Feedback Surveys 2006 to 2016

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*	Overall by SAM
Overall by year	89%	91%	93%	79%	74%	64%	56%	67%	68%	74%	73%	71%
Diabetes	91%	92%	94%	79%	73%	65%	59%	68%	70%	71%	76%	73%
Hypertension	90%	91%	93%	79%	74%	64%	60%	67%	70%	74%	77%	73%
Asthma	90%	93%	94%	81%	74%	67%	64%	72%	71%	77%	74%	77%
Coronary artery disease	89%	90%	92%	80%	74%	67%	59%	67%	68%	72%	70%	73%
Depression	88%	91%	93%	78%	72%	64%	57%	67%	64%	71%	68%	73%
Heart failure	87%	90%	91%	78%	74%	66%	58%	64%	67%	74%	71%	72%
Pain management		92%	92%	78%	75%	64%	55%	67%	71%	76%	75%	72%
Well child care		92%	91%	80%	74%	65%	56%	63%	66%	75%	73%	71%
Health behavior			93%	77%	74%	63%	56%	63%	64%	72%	68%	70%
Maternity care			93%	80%	77%	67%	58%	62%	64%	72%	71%	71%
Care of vulnerable elders				78%	74%	64%	56%	67%	71%	76%	72%	69%
Childhood illness				79%	76%	64%	56%	62%	65%	72%	72%	68%
Cerebrovascular disease					76%	63%	55%	69%	67%	73%	71%	66%
Preventive care						56%	49%	69%	67%	74%	74%	59%
Hospital medicine							57%	67%	71%	76%	74%	71%
Mental health							52%	65%	65%	73%	65%	63%

Data from 2006 to 2012 reported by Brooks et al¹⁸.

*Data from January 2016 to July 2016.

Table 2. Ratings of Content Appropriateness of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Knowledge Assessment (KA) from 2006 to 2016

Module Name	Content Is Appropriate for My Practice, Mean (95% CI)										
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*
Diabetes	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)
Hypertension	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.2 (5.2–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)
Asthma	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.1 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.1)
Coronary artery disease	4.8 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.9–5.0)	4.8 (4.8–4.9)	4.8 (4.8–4.8)	4.8 (4.8–4.8)	4.8 (4.8–4.9)	4.8 (4.7–4.8)	4.7 (4.7–4.8)	4.6 (4.6–4.7)	4.7 (4.6–4.7)
Depression	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	4.9 (4.8–4.9)	5.0 (4.9–5.0)
Heart failure	5.1 (5.0–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	4.9 (4.9–5.0)
Pain management†		5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.2 (5.2–5.2)	5.2 (5.1–5.2)	5.2 (5.2–5.3)	5.2 (5.1–5.2)	5.2 (5.2–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)
Well child care†		4.9 (4.9–5.0)	4.8 (4.8–4.9)	4.9 (4.8–5.0)	5.2 (5.1–5.2)	5.3 (5.2–5.3)	5.2 (5.2–5.2)	5.2 (5.2–5.3)	5.2 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.2)
Health behavior†			5.1 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.0 (4.9–5.1)
Maternity care†			4.8 (4.8–4.9)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.7 (4.6–4.8)	4.8 (4.7–4.9)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (4.9–5.1)
Care of vulnerable elders				5.0 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)
Childhood illness				4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.8 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.8–4.9)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.8–4.9)
Cerebrovascular disease					4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.8–4.9)
Preventive care						5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)
Hospital medicine							4.9 (4.8–5.1)	4.9 (4.9–5.0)	4.8 (4.8–4.9)	4.7 (4.7–4.8)	4.8 (4.7–4.8)
Mental health							4.9 (4.8–4.9)	4.8 (4.7–4.9)	4.8 (4.7–4.9)	4.7 (4.6–4.8)	4.8 (4.7–4.9)

†The ratings are on a scale from 1 (strongly disagree) to 6 (strongly agree) with 95% confidential interval (CI) in parentheses.

*Data from January 2016 to July 2016.

†The ratings of these 4 SAMs were rescaled from 1 to 4 to 1 to 6.

Table 3. Ratings of Content Appropriateness of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Knowledge Assessment (KA) from 2006 to 2016

Module name	Content Was Presented at an Appropriate Level, Mean (95% CI)										
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*
Diabetes	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)
Hypertension	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)
Asthma	5.1 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)
Coronary artery disease	4.8 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.9–5.0)	4.9 (4.9–4.9)	4.9 (4.9–4.9)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–4.9)	4.9 (4.8–4.9)	4.8 (4.7–4.8)	4.8 (4.7–4.8)
Depression	5.0 (4.9–5.0)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (4.9–5.0)	4.9 (4.8–4.9)	4.9 (4.9–5.0)
Heart failure	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.0 (5.0–5.1)	4.9 (4.9–5.0)
Pain management†		5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.2–5.2)	5.2 (5.2–5.3)	5.1 (5.1–5.2)	5.2 (5.2–5.3)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)
Well child care†		4.8 (4.8–4.9)	4.8 (4.8–4.9)	4.8 (4.8–4.9)	5.2 (5.1–5.2)	5.2 (5.2–5.3)	5.3 (5.2–5.3)	5.2 (5.2–5.3)	5.2 (5.2–5.3)	5.2 (5.2–5.2)	5.2 (5.1–5.3)
Health behavior†			5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.2–5.2)	5.2 (5.2–5.2)	5.2 (5.2–5.3)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.0)
Maternity care†			5.0 (4.9–5.0)	5.1 (5.0–5.1)	5.2 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.1 (5.1–5.0)
Care of vulnerable elders			5.0 (4.9–5.0)	5.1 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)
Childhood illness				5.0 (5.0–5.1)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)
Cerebrovascular disease				5.0 (4.9–5.0)	5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)
Preventive care					5.1 (5.1–5.1)	5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)
Hospital medicine						5.1 (5.1–5.1)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.8 (4.7–4.9)
Mental health							4.9 (4.8–5.0)	4.9 (4.8–5.0)	4.8 (4.7–4.9)	4.7 (4.6–4.8)	4.8 (4.7–4.9)

The ratings are on a scale from 1 (strongly disagree) to 6 (strongly agree) with 95% confidential interval (CI) in parentheses.

*Data from January 2016 to July 2016.

†The ratings of these 4 SAMs were rescaled from 1 to 4 to 1 to 6.

Table 4. Ratings of Content Appropriateness of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Clinical Simulation (CS) from 2006 to 2016

Module Name	Content Is Appropriate for My Practice, Mean (95% CI)										
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*
Diabetes	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)
Hypertension	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)
Asthma	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.8 (4.8–4.9)
Coronary artery disease	4.8 (4.7–4.8)	4.8 (4.8–4.9)	4.9 (4.9–4.9)	4.9 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.9 (4.8–4.9)
Depression	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.9 (4.8–5.0)
Heart failure	5.0 (4.9–5.1)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (4.9–5.0)	4.9 (4.8–5.0)
Pain management†		4.9 (4.9–5.0)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.0 (5.0–5.1)	5.0 (4.9–5.0)	4.9 (4.8–5.0)
Well child care†		4.7 (4.7–4.8)	4.8 (4.7–4.8)	4.7 (4.6–4.8)	4.9 (4.9–5.0)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.1)
Health behavior†			5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.1 (5.0–5.1)	5.0 (4.9–5.1)	5.1 (5.0–5.2)	5.0 (4.9–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	4.8 (4.7–4.9)
Maternity care†			4.7 (4.6–4.8)	4.6 (4.5–4.7)	4.6 (4.5–4.7)	4.6 (4.5–4.7)	4.5 (4.4–4.6)	4.5 (4.4–4.6)	4.6 (4.5–4.7)	4.7 (4.6–4.7)	4.6 (4.5–4.8)
Care of vulnerable elders				4.8 (4.8–4.9)	4.9 (4.8–4.9)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.8 (4.8–4.9)	4.8 (4.7–4.9)
Childhood illness				4.7 (4.6–4.7)	4.7 (4.6–4.7)	4.7 (4.6–4.7)	4.7 (4.6–4.8)	4.8 (4.7–4.8)	4.7 (4.7–4.8)	4.7 (4.7–4.8)	4.7 (4.6–4.8)
Cerebrovascular disease					4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.8 (4.7–4.9)	4.8 (4.7–4.9)
Preventive care						5.1 (5.1–5.1)	5.0 (5.0–5.1)	(5.0–5.1)	5.0 (5.0–5.1)	4.9 (4.9–5.0)	4.9 (4.8–5.0)
Hospital medicine							5.1 (5.1–5.1)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.8 (4.7–4.8)	4.7 (4.7–4.8)
Mental health							4.9 (4.8–5.0)	4.8 (4.7–4.9)	4.8 (4.7–4.9)	4.7 (4.6–4.8)	4.8 (4.6–4.9)

The ratings are on a scale from 1 (strongly disagree) to 6 (strongly agree) with 95% confidential interval (CI) in parentheses.

*Data from January 2016 to July 2016.

†The ratings of these 4 SAMs were rescaled from 1 to 4 to 1 to 6.

Table 5. Ratings of Content Appropriateness of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Clinical Simulation (CS) from 2006 to 2016

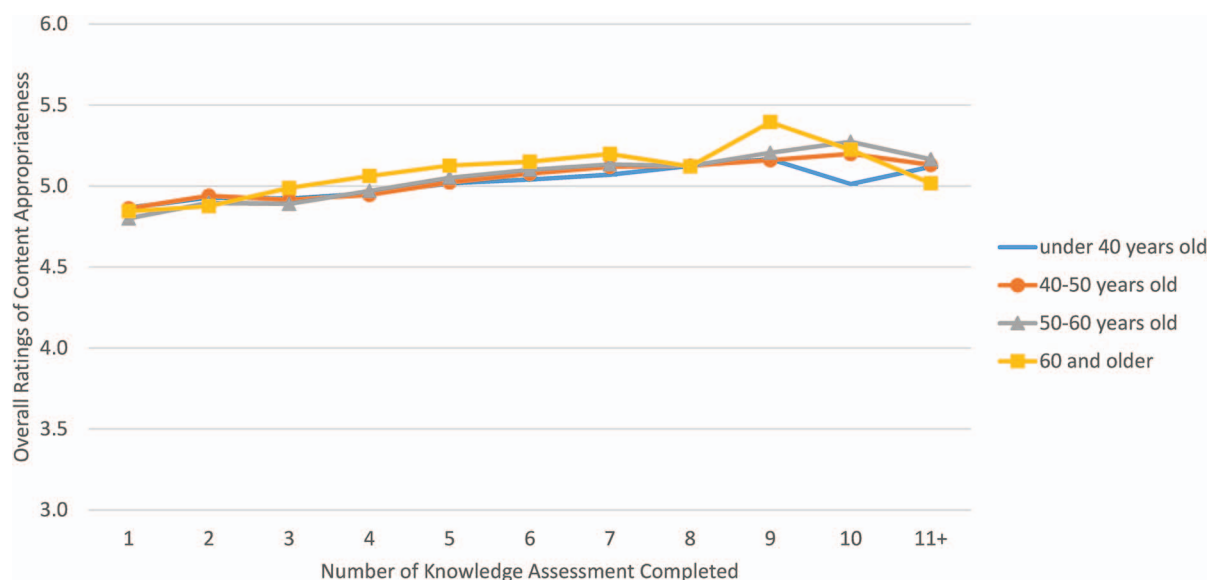
Module Name	Content Was Presented at an Appropriate Level, Mean (95% CI)										
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016*
Diabetes	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–5.0)
Hypertension	4.9 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)
Asthma	5.0 (4.9–5.0)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	4.9 (4.8–4.9)
Coronary artery disease	4.8 (4.7–4.8)	4.8 (4.8–4.9)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–5.0)
Depression	4.9 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.8 (4.8–4.9)
Heart failure	5.0 (4.9–5.0)	5.0 (5.0–5.1)	5.1 (5.1–5.2)	5.0 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.0–5.1)	5.0 (4.9–5.1)	4.9 (4.9–5.0)
Pain management†	4.9 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.1 (5.0–5.1)	5.0 (5.0–5.1)	5.0 (4.9–5.0)	4.8 (4.8–4.9)
Well child care†		4.6 (4.6–4.7)	4.7 (4.6–4.7)	4.6 (4.5–4.7)	4.7 (4.6–4.7)	4.8 (4.7–4.8)	4.9 (4.9–5.0)	4.9 (4.8–5.0)	4.9 (4.8–4.9)	5.1 (5.0–5.1)	5.1 (5.0–5.1)
Health behavior†			5.1 (5.0–5.2)	5.1 (5.0–5.1)	5.1 (5.1–5.2)	5.1 (5.1–5.2)	5.2 (5.1–5.2)	5.1 (5.1–5.2)	5.1 (5.0–5.2)	5.1 (5.1–5.2)	5.0 (4.9–5.1)
Maternity care†			4.7 (4.6–4.8)	4.5 (4.5–4.6)	4.5 (4.4–4.6)	4.5 (4.4–4.6)	4.6 (4.5–4.7)	4.6 (4.5–4.7)	4.6 (4.6–4.7)	4.6 (4.5–4.7)	4.6 (4.5–4.7)
Care of vulnerable elders				4.8 (4.7–4.8)	4.8 (4.8–4.9)	4.9 (4.8–4.9)	4.9 (4.8–4.9)	4.8 (4.8–4.9)	4.8 (4.7–4.8)	4.8 (4.8–4.9)	4.8 (4.7–4.9)
Childhood illness				4.5 (4.4–4.5)	4.5 (4.5–4.6)	4.5 (4.5–4.6)	4.5 (4.5–4.6)	4.6 (4.6–4.7)	4.6 (4.6–4.7)	4.6 (4.6–4.7)	4.6 (4.5–4.7)
Cerebrovascular disease					5.0 (4.9–5.0)	4.9 (4.9–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	4.9 (4.9–5.0)	4.9 (4.8–4.9)	4.8 (4.7–4.9)
Preventive care						5.1 (5.1–5.2)	5.0 (5.0–5.0)	5.0 (5.0–5.0)	5.0 (5.0–5.1)	5.0 (4.9–5.0)	4.9 (4.8–5.0)
Hospital medicine							4.8 (4.7–5.0)	4.8 (4.8–4.9)	4.9 (4.8–4.9)	4.7 (4.7–4.8)	4.7 (4.6–4.8)
Mental health							4.9 (4.8–5.0)	4.9 (4.8–5.0)	4.8 (4.8–4.9)	4.7 (4.6–4.8)	4.8 (4.7–4.9)

The ratings are on a scale from 1 (strongly disagree) to 6 (strongly agree) with 95% confidential interval (CI) in parentheses.

*Data from January 2016 to July 2016.

†The ratings of these 4 SAMs were rescaled from 1 to 4 to 1 to 6.

Figure 1. Overall ratings of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Knowledge Assessment (KA) content appropriateness by increasing numbers of modules completed and by age at first module completion.



Childhood Illness CS's, whose content appropriateness was rated approximately one half point lower than their KA counterparts.

Figure 1 illustrates differences in FPs' opinions about overall content appropriateness by charting mean ratings on a continuum of increasing number of KAs completed, by age at first KA. Regardless of the topic area, we observed that the ratings were lowest among FPs who had completed 1 KA but increased as they completed more KAs over time (2006 to 2016). This trend applied to both younger and older FPs. Nevertheless, the ratings appeared to be higher among older FPs and lower among younger FPs. The corresponding trends of CSs closely mirrored those of KA (Figure 2). Nonetheless, ratings of content appropriateness of CS were lower than those of the KAs.

Discussion

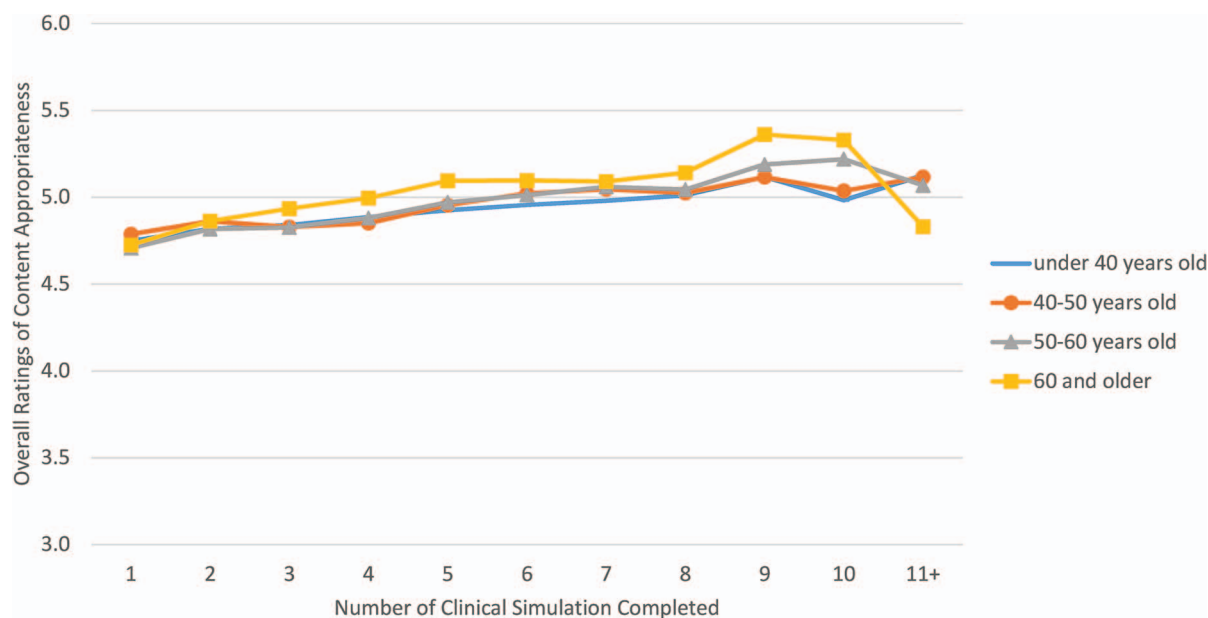
Building on a decade's worth of Diplomate feedback, this study provides evidence that FPs generally value ABFM SAMs. Contrary to claims that MOC activities are irrelevant and unnecessary,⁹⁻¹¹ the content of all 16 SAM topics was regarded by most FPs to be appropriate for practice and presented at an appropriate level. The fairly stable ratings over the life cycle of each individual KA and CS further suggested that the ABFM's process for updating both components kept the content rele-

vant to FPs' practice.²¹ As we found no evidence of large differences between responders and nonresponders and nearly 70% of completed SAMs had feedback surveys, these findings are representative of a majority of ABFM Diplomates.

Compared with FPs who completed fewer KAs or CSs, those who completed more activities tended to rate the content higher. That is, FPs' opinions about the content became more favorable as they engaged with more activities. Ensuring that the content continues to be relevant for practice is crucial for the success of MOC Part II activities, which are intended to serve as a lifelong learning platform for physicians. The consistently high ratings of content appropriateness over the past decade suggested that ABFM SAMs served their purpose well.

The lowest ratings reflected the opinions of those who completed fewer SAMs and were likely FPs who were newly introduced to MOC. This may imply that younger cohorts of FPs had different initial experiences with or expectations of SAMs specifically, and perhaps MOC in general. Therefore, listening to and addressing the concerns of physicians who just began their MOC journey may help boost participation and satisfaction over the long term. This finding also highlights the importance of continuing to monitor Diplomates' feedback. As a result of an analysis of these feedback

Figure 2. Overall ratings of American Board of Family Medicine (ABFM) Self-Assessment Module (SAM) Clinical Simulation (CS) content appropriateness by increasing numbers of modules completed and by age at first module completion.



survey data showing that most critiques of the SAM were associated with technical issues in the CS,^{18,22} in July 2016, the ABFM split the SAM into 2 independent activities: the KSA and the CSA. The function of soliciting FP feedback remains and will continue to assist the ABFM in evaluating the performance of its new KSA and CSA modules.¹⁹

Limitations

This study has several limitations. First, the rating scales were not uniform across all SAM topics. Although we aligned different scales to be comparable, the weighted mean ratings may have skewed physicians' opinions about SAM content that was originally rated on a scale of 1 to 4. Nevertheless, the rating patterns of these SAMs agreed with those SAMs on a scale of 1 to 6. Second, survey respondents tended to be older than nonrespondents which suggests a potential overestimation of the results as older FPs viewed SAMs more favorably than younger FPs. However, the age differences were small despite being statistically significant and would be unlikely to affect the main results in a significant way.

Conclusion

This study provides substantial evidence supporting the perceived value of ABFM SAMs to Diplo-

mates. Over the past decade, Diplomates rated the content of ABFM SAMs as appropriate and presented at an appropriate level across 16 diverse clinical areas. Importantly, FPs who completed more SAMs tended to have more favorable opinions of the content than those who completed fewer. Continued monitoring of feedback will play a critical role in keeping the content of MOC programs relevant for physicians' practice.

To see this article online, please go to: <http://jabfm.org/content/32/1/79.full>.

References

1. American Board of Medical Specialties. Standards for the ABMS Program for Maintenance of Certification (MOC). 2014. Available from: <https://www.abms.org/media/1109/standards-for-the-abms-program-for-moc-final.pdf>.
2. Lawrence HC III. Letter to Dr. Nora. 2017. <http://www.medtees.com/content/SpSocMOCLetter.pdf>. Accessed June 14, 2018.
3. Peterson LE, Blackburn BE, Puffer JC, Phillips RL Jr. Family physicians' quality interventions and performance improvement through the ABFM diabetes performance in practice module. *Ann Fam Med* 2014;12:17–20.
4. Peterson LE, Blackburn B, Puffer JC, Phillips RL Jr. Family Physicians' Quality Interventions and Performance Improvement for Hypertension through

- Maintenance of Certification. *J Healthc Qual* 2016; 38:175–86.
5. Peterson LE, Blackburn B, Phillips RL, Puffer JC. Improving quality of care for diabetes through a maintenance of certification activity: family physicians' use of the chronic care model. *J Contin Educ Health Prof* 2014;34:47–55.
6. Galliher JM, Manning BK, Petterson SM, et al. Do professional development programs for Maintenance of Certification (MOC) affect quality of patient care? *J Am Board Fam Med* 2014;27:19–25.
7. Elward K, Blackburn B, Peterson LE, Greenawald M, Hagen MD. Improving quality of care and guideline adherence for asthma through a group self-assessment module. *J Am Board Fam Med* 2014;27:391–8.
8. American Board of Medical Specialty. MOC participation improves patient care, leads to advocacy efforts and practice transformation. 2015. Available from: <http://www.abms.org/news-events/moc-participation-improves-patient-care-leads-to-advocacy-efforts-and-practice-transformation/>.
9. Lipner RS, Hess BJ, Phillips RL Jr. Specialty board certification in the United States: Issues and evidence. *J Contin Educ Health Prof* 2013;33 Suppl 1:S20–S35.
10. Hawkins RE, Lipner RS, Ham HP, Wagner R, Holmboe ES. American Board of Medical Specialties Maintenance of Certification: Theory and evidence regarding the current framework. *J Contin Educ Health Prof* 2013;33 Suppl 1:S7–S19.
11. Cook DA, Holmboe ES, Sorensen KJ, Berger RA, Wilkinson JM. Getting maintenance of certification to work: A grounded theory study of physicians' perceptions. *JAMA Intern Med* 2015;175:35–42.
12. Lipner RS, Bylsma WH, Arnold GK, Fortna GS, Tooker J, Cassel CK. Who is maintaining certification in internal medicine—And why? A national survey 10 years after initial certification. *Ann Intern Med* 2006;144:29–36.
13. Hagen MD, Ivins DJ, Puffer JC, et al. Maintenance of certification for family physicians (MC-FP) self assessment modules (SAMs): the first year. *J Am Board Fam Med* 2006;19:398–403.
14. Holmboe ES, Lipner R, Greiner A. Assessing quality of care: Knowledge matters. *JAMA* 2008; 299:338–40.
15. Peterson LE, Eden A, Cochrane A, Hagen M. Physician satisfaction with and practice changes resulting from American Board of Family Medicine Maintenance of Certification performance in practice modules. *J Contin Educ Health Prof* 2016;36:55–60.
16. Ackerman KG, Lee B, Kushner JA. Dissatisfaction with maintenance of certification in academic pediatrics. *Pediatr Res* 2016;79:240–2.
17. Teirstein PS. Boarded to death—Why maintenance of certification is bad for doctors and patients. *N Engl J Med* 2015;372:106–8.
18. Brooks EM, Gonzalez M, Eden AR, O'Neal J, Sabo RT, Etz RS. What Family Physicians really think of Maintenance of Certification Part II activities. *J Contin Educ Health Prof* 2017;37:223–9.
19. American Board of Family Medicine. KSA & CSA feedback survey: Results six weeks post-SAM split. Phoenix 2016;7. Available from: <https://www.theabfm.org/about/newsletter2016december.pdf>.
20. American Board of Family Medicine. Self-Assessment Module (SAM) separated into two activities. Phoenix 2016;3. Available from: <https://www.theabfm.org/about/newsletter.aspx>.
21. Hagen MD, Quan MA, Fain R, DeBord W. ABFM's self-assessment module (SAM) revision process. *Ann Fam Med* 2015;13:88–9.
22. Puffer JC. A message from the president of the American Board of Family Medicine (ABFM). *J Am Board Fam Med* 2016;29:815–6.