

# Physician Scores on a National Clinical Skills Examination as Predictors of Complaints to Medical Regulatory Authorities

Robyn Tamblyn, PhD

Michal Abrahamowicz, PhD

Dale Dauphinee, MD

Elizabeth Wenghofer, PhD

André Jacques, MD

Daniel Klass, MD

Sydney Smee, MSc

David Blackmore, PhD

Nancy Winslade, PharmD

Nadyne Girard, MSc

Roxane Du Berger, MSc

Ilona Bartman, MA

David L. Buckeridge, MD, PhD

James A. Hanley, PhD

**D**ECADES OF RESEARCH HAVE confirmed that poor skills in patient communication are associated with lower levels of patient satisfaction, higher rates of complaints, an increased risk of malpractice claims, and poorer health outcomes.<sup>1-16</sup> Medical schools have responded by incorporating training in patient communication and clinical skills into the curriculum. However, these skills were not systematically evaluated, nor was a minimum level of proficiency required for medical licensure.<sup>17</sup> To address this problem, licensure reforms were undertaken in North America.<sup>18</sup> The Medical Council of Canada (MCC) (1993),<sup>19</sup> the Educational Commission for Foreign

**Context** Poor patient-physician communication increases the risk of patient complaints and malpractice claims. To address this problem, licensure assessment has been reformed in Canada and the United States, including a national standardized assessment of patient-physician communication and clinical history taking and examination skills.

**Objective** To assess whether patient-physician communication examination scores in the clinical skills examination predicted future complaints in medical practice.

**Design, Setting, and Participants** Cohort study of all 3424 physicians taking the Medical Council of Canada clinical skills examination between 1993 and 1996 who were licensed to practice in Ontario and/or Quebec. Participants were followed up until 2005, including the first 2 to 12 years of practice.

**Main Outcome Measure** Patient complaints against study physicians that were filed with medical regulatory authorities in Ontario or Quebec and retained after investigation. Multivariate Poisson regression was used to estimate the relationship between complaint rate and scores on the clinical skills examination and traditional written examination. Scores are based on a standardized mean (SD) of 500 (100).

**Results** Overall, 1116 complaints were filed for 3424 physicians, and 696 complaints were retained after investigation. Of the physicians, 17.1% had at least 1 retained complaint, of which 81.9% were for communication or quality-of-care problems. Patient-physician communication scores for study physicians ranged from 31 to 723 (mean [SD], 510.9 [91.1]). A 2-SD decrease in communication score was associated with 1.17 more retained complaints per 100 physicians per year (relative risk [RR], 1.38; 95% confidence interval [CI], 1.18-1.61) and 1.20 more communication complaints per 100 practice-years (RR, 1.43; 95% CI, 1.15-1.77). After adjusting for the predictive ability of the clinical decision-making score in the traditional written examination, the patient-physician communication score in the clinical skills examination remained significantly predictive of retained complaints (likelihood ratio test,  $P < .001$ ), with scores in the bottom quartile explaining an additional 9.2% (95% CI, 4.7%-13.1%) of complaints.

**Conclusion** Scores achieved in patient-physician communication and clinical decision making on a national licensing examination predicted complaints to medical regulatory authorities.

*JAMA.* 2007;298(9):993-1001

www.jama.com

Medical Graduates (1998),<sup>20</sup> and most recently the United States Medical Licensing Examination (USMLE) (2004)<sup>21</sup> have all introduced a clinical skills examination (CSE)—a nationally standardized assessment of patient-physician communication, clinical history taking, and examination skills—as a requirement for licensure. All US and Canadian medical

**Author Affiliations:** Departments of Medicine (Drs Tamblyn and Dauphinee) and Epidemiology & Biostatistics (Drs Abrahamowicz, Winslade, Buckeridge, and Hanley, and Mss Girard and Du Berger), McGill University, Montreal, Quebec, Canada; Ontario College of Physicians and Surgeons, Toronto, Ontario, Canada (Drs Wenghofer and Klass); Quebec College of Physicians, Montreal (Dr Jacques); and Medical Council of Canada, Ottawa, Ontario (Dr Blackmore and Mss Smee and Bartman).

**Corresponding Author:** Robyn Tamblyn, PhD, McGill University, 1140 Pine Ave W, Montreal, QC Canada, H3A 1A3 (robyn.tamblyn@mcgill.ca).

For editorial comment see p 1057.



CME available online at  
www.jama.com

school graduates must now pass a multiple-case standardized patient assessment, where patient and physician examiners observe and grade clinical and communication skills to predict a candidate's competence to practice.

While mandatory assessment of clinical and communication skills is supported by the general public,<sup>22</sup> concerns have been raised about the cost of the examination and the lack of evidence that a 1-day assessment could predict future practice, particularly as it relates to deficiencies in patient-physician communication.<sup>23-27</sup>

Since instituting its CSE for all Canadian physicians, the MCC has tested more than 25 000 medical graduates using an examination format similar to the USMLE Step 2 clinical skills examination.<sup>19</sup> We investigated the ability of CSEs to predict future complaints in medical practice. We tested the hypothesis that lower scores in patient-physician communication would be associated with a higher rate of patients' complaints about quality of care and communication. We also assessed whether the use of clinical examination scores improved the prediction of complaints beyond results from the traditional written examination.

## METHODS

### Context

In Canada and the United States, medical regulatory authorities (state medical boards and provincial colleges of physicians and surgeons) use a common framework to govern how physicians are trained, accepted into practice, regulated, disciplined, and removed from practice.<sup>28-32</sup> A principal obligation of state and provincial medical regulatory authorities in both countries is to address and resolve public complaints against physicians. In accordance with a common set of principles and procedures, all complaints that are received in writing are investigated. A triage system is used to collect information from the patient and physician for each complaint, weed out frivolous or vexatious actions, and undertake informal steps

to attain early resolution of minor issues. When these informal steps are either unsuccessful or deemed inappropriate, the complaint is managed by a more formal committee or panel process that determines further action. Most complaints are resolved through a graded series of regulatory actions, typically education, cautions, and warnings. For the most serious complaints, and for all complaints involving issues of sexual misconduct, formal disciplinary hearings of a quasi-judicial nature are convened. These hearings can result in a variety of sanctions up to loss of license. When a patient complaint about a physician is made directly to a hospital, the hospital in most state and provincial jurisdictions is required to report problems of professional misconduct to the medical regulatory authority.

### Design and Population

The cohorts of physicians who took the MCC clinical skills examination between 1993 and 1996 and were licensed to practice in Ontario and/or Quebec were identified. Nearly two-thirds of the Canadian population and approximately 50% of all physicians reside in these 2 provinces. All complaints filed against these physicians with the medical regulatory authority in either province were retrieved between the date of licensure and March 2005. The MCC identified the 6677 physicians taking the examination during this period and provided the first and last name, sex, medical school, and year of graduation of each candidate to the medical regulatory authority in Ontario and Quebec. These 5 nominal fields were used to link to the registry of licensed physicians in each province. Physicians who matched on all fields were retained. Partial matches were manually inspected and adjudicated. Specialty, postgraduate training location and dates, and license year were obtained from the provincial medical regulatory files as well as from the national training registry of all physicians completing postgraduate medical training in Canada. Of the 6677 phy-

sicians, 8.6% could not be linked to Ontario/Quebec medical regulatory files or the national postgraduate training registry. Compared with linked physicians, unlinked physicians were more likely to be older (>45 years, 44.2% vs 11.4%;  $\chi^2 P < .001$ ), men (73.6% vs 57.4%;  $\chi^2 P < .001$ ), have trained outside Canada (83.4% vs 12.7%;  $\chi^2 P < .001$ ), have not yet passed the CSE (15.7% vs 1.8%;  $\chi^2 P < .001$ ), and have lower traditional written examination scores (495.4 vs 524.7; *t* test  $P < .001$ ) and CSE scores (436.8 vs 517.8; *t* test  $P < .001$ ).

Physician identity and confidential information were protected by replacing all nominal data with an MCC-generated study number, which was used to link demographic, score, and complaint files for each study physician. The McGill Faculty of Medicine institutional review board provided ethical approval. The provincial privacy commission, the Ontario and Quebec medical regulatory authorities, and MCC approved and oversaw data access, linkage, and anonymization procedures.

### Measurement of Complaints

Provincial medical regulatory authorities collect standardized information for each written complaint against a physician. This information includes the names of the patients and physicians involved, and a description of the problem, circumstances, medical interventions, outcome, and the location of the incident. The investigation process includes a review of the letter with the complainant, the physician response, the patient's medical records, information from the hospital if applicable (eg, for surgical complications), and information from witnesses. All evidence is reviewed by physician investigators (Quebec) or a complaints committee (Ontario) who determine the legitimacy of the complaint, the type and seriousness of problem, and the recommended approach for resolution and subsequent action. Complaints are classified by investigators into 1 of 55 (Quebec) or 57 (Ontario) mutually exclu-

sive categories (eg, complication due to medical or surgical error, breach of confidentiality, incomplete medical reports), along with the outcome (retained or not) and the action (warning, counseling/training, license withdrawal, suspension, or restriction).

All complaints recorded for study physicians were retrieved by medical regulatory personnel. Data included the physician study number, date of filing and closure, the classification of problem type, and the outcome (retention decision and action taken). Complaint classification codes from the respective regulatory authorities were grouped into 6 categories based on comparable groupings used by the Ontario and Quebec regulatory authorities: communication and attitude; quality of care; professionalism; office-related problem; physician health-related behavior problem (eg, mental illness); and other (eg, false advertising). Assignment of complaint classification codes was independently verified by medical regulatory investigators who arbitrated disagreements on final assignment.

The primary outcome was the complaint rate: the number of complaints retained as valid by the medical regulatory authority after investigation per year of practice time. Because judgment about the validity of a complaint may vary between provincial regulatory authorities, we conducted a sensitivity analysis including all complaints to assess whether our findings were influenced by retention decisions. The subset of retained complaints that were related to problems in communication and quality of care were assessed as secondary outcomes, as these problems should be more strongly associated with the competencies being assessed by the examination.

The complaint rate for each physician was calculated using as the denominator years in practice, defined as the number of years between the final year of postgraduate training exit date and the end of follow-up (March 2005). To assess the validity of using exit date from postgraduate training as the start-

ing date for practice time, we retrieved for 1161 Quebec physicians a count of the number of years in which the physician billed for patient services to the provincial insurance agency. In comparison with actual billing data between 1993 and 2003, our approach modestly overestimates the number of practice-years (mean [SD] from billing, 4.2 [2.4] years; from training exit year estimate, 4.9 [2.2] years). However, there was very good agreement between the 2 methods (intra-class correlation, 0.67; 95% confidence interval [CI], 0.54-0.75) and no relationship between practice-years and communication score (Pearson  $r=-0.06$ ). Thus, potential errors in measurement of practice-years should not confound the association between complaints and communication score.

#### Medical Council of Canada Examinations

**Traditional Written Examination.** This examination tests an individual's competence to enter postgraduate training. It is generally taken at the end of medical school and must be passed to be eligible for licensure. Medical knowledge is assessed using approximately 450 multiple-choice questions to assess knowledge in medicine, surgery, obstetrics-gynecology, psychiatry, pediatrics, and preventive medicine.<sup>33</sup> Clinical decision making is assessed using key feature problems.<sup>34</sup> Examinees are asked to respond to critical aspects of diagnosis or management in 36 to 40 clinical problems using write-in or menu-selection response formats.<sup>34</sup> Unlike multiple-choice questions, key feature questions focus exclusively on the components of a case where physicians are required to make critical decisions where errors could have an effect on patient outcome. Grading is based on the relative quality of the response, rather than a single correct answer, and errors of both omission and commission are considered in scoring. The score is calculated as the weighted sum of the multiple-choice (weight=0.75) and clinical decision-making skills components (weight=0.25), where the

weights reflect the amount of testing time devoted to each component. A criterion-based passing score is established by a modified Nedelsky method,<sup>33,34</sup> and scores for first-time takers are standardized to a mean (SD) of 500 (100). For the study population, the Cronbach  $\alpha$  estimate of the reliability of the written examination varied from 0.90 to 0.92 for the multiple-choice component, and from 0.60 to 0.69 for the clinical decision-making component in different administrations.

**Clinical Skills Examination.** This examination tests competence in data collection (history, physical examination), patient communication, and problem solving (diagnosis and management) through a 20-case objective structured clinical examination, and can be taken after 1 year of postgraduate training.<sup>19</sup> Most physicians take the examination in the second postgraduate year or the first half of the third postgraduate year (93% of physicians taking the examination between 1993 and 1996). Data collection is assessed in a 5- or 10-minute interaction with a standardized patient, by trained physician observers using case-specific checklists.<sup>19</sup> Patient-centered communication is assessed in 3 to 4 cases, selected to represent situations where communication is required for effective management (eg, discuss refusal of treatment for a terminal illness, counsel an adolescent about birth control). Examples of patient-physician communication that would receive a low score include condescending, offensive, or judgmental behaviors, or ignoring patient responses during the encounter. Problem solving is assessed by post-counter written responses to short-answer questions on diagnosis, investigation, interpretation of test results, and management. Responses are scored by physician examiners using an answer key. The passing score for the overall examination is established using criterion-referenced methods,<sup>19,33-35</sup> and scores for first-time takers are standardized to a mean (SD) of 500 (100). For the study population, the Cronbach  $\alpha$  estimate of

the reliability of the CSE scores ranged from 0.25 to 0.50 for communication, 0.59 to 0.75 for data acquisition, and 0.41 to 0.67 for problem solving in different administrations.

### Covariates

Physician characteristics that may be associated with communication ability or complaint rate were measured as potential confounders and effect modifiers.<sup>6,10</sup> They included information on the sex of the physician, international medical graduate status, and specialty, which were retrieved from the MCC master file, postgraduate training registry, and the medical regulatory authorities. Practice province also was considered a potential confounder because differences may exist in health service delivery and the management of complaints between jurisdictions.

### Statistical Analysis

Correlations between examination scores were estimated by Pearson product-moment correlation coefficients. Score reliability was assessed using a weighted Cronbach  $\alpha$ , where weights were based on the number of candidates taking the examination in each administration. Disattenuated correlations also were calculated to determine the expected correlation if both scores were measured with perfect reliability, using the formula<sup>36</sup>

$$r_{xy} \text{ Disattenuated} = \frac{r_{xy}}{\sqrt{\text{Reliability}_x \times \text{Reliability}_y}}$$

The relationship between the CSE scores and complaint rate was assessed using multivariate Poisson regression (SAS version 9.1, SAS Institute, Cary, North Carolina), adjusting for physician sex, specialty, country of training (Canada or international), and province. A 2-sided test with a *P* value of .05 was used to assess statistical significance. Number of complaints was the dependent variable, and number of years in practice was used to measure person-time for each physician. The predictive ability of each examination score was assessed in a separate model that adjusted for sex, specialty, inter-

national medical graduate status, and practice province, using continuous scores as well as score quartiles. To determine if the relationship between examination scores and complaints was modified by characteristics that may be associated with communication scores, including practice jurisdiction, physician sex, and foreign training, we assessed interactions between the examination score and these characteristics and used the likelihood ratio test to determine if the interaction terms improved the model fit.

Licensing examinations aim to assess a required level of proficiency, and thus minimum thresholds of communication ability may exist, below which the complaint rate is high and above which the rate is lower and relatively uniform. To assess whether a linear association provided an appropriate representation of the association between examination score and the complaint rate, we tested the multivariate Poisson models for non-linearity using generalized additive models (GAM) nonparametric extension of Poisson regression.<sup>37</sup> The adjusted effect of examination score was estimated using smoothing splines with 4 *df* and the statistical significance of the nonlinear effect was tested by nonparametric  $\chi^2$  test. All models were estimated separately for primary and secondary outcomes.

To determine if including the CSE communication score improved the prediction of complaints beyond the traditional written examination results, we first estimated the independent relationship between scores achieved in the traditional written examination and complaint rate. The CSE communication score was then added to the model that included the traditional written examination score, and improvement in the prediction of complaints was assessed by the likelihood ratio test. The explanatory power of the CSE communication score in predicting complaints was estimated by the population attributable fraction, the proportion of all complaints that were explained by physicians in the bottom communication score quartile,<sup>38</sup> after adjustment for existing predictors.

Power was estimated using the approach proposed by Signorini<sup>39</sup> for Poisson regression. Based on a type I error of 5%, a baseline complaint rate of 3.1% in the study population, and 3424 physicians followed up for a mean 6.5 years, the study had a power of 95% to detect a relative rate difference of 12% per 2-SD decrease in score.

### RESULTS

Among 6677 physicians taking the CSE between 1993 and 1996, 3424 (51.3%) were licensed to practice in Ontario and/or Quebec. At the time of the examination, 71.6% of study physicians were 25 to 30 years of age, 55.5% were men, and 12.3% were international medical graduates. Following the examination, 84% completed postgraduate training in primary care or medical subspecialties, and two-thirds entered practice in Ontario (TABLE 1). The mean score of the study population for both the clinical skills and traditional written examinations was approximately one-quarter of an SD above 500. However, the range was considerable—approximately 7 SDs for the CSE and 5 SDs for the traditional written examination. Overall, 230 physicians (6.7%) failed the CSE on the first attempt, and 52 of these physicians never passed the CSE but were licensed to practice during the transition to the new licensure requirements.

Correlations between the clinical skills and traditional written examinations overall scores and subscores varied between  $r=0.10$  and  $r=0.40$  (TABLE 2). The communication score had the lowest correlation with the traditional written examination scores and with other scores on the CSE. Even when corrected for unreliability, the correlation between the communication and traditional written examination scores was low (disattenuated  $r=0.23$ ). Communication ability previously has been shown to be a domain independent from more cognitive abilities that are assessed in traditional written examinations.<sup>40</sup>

Overall, 1116 complaints were filed in a total of 22 585 practice-years (4.9

complaints per 100 practice-years) (TABLE 3). The mean (SD) follow-up time per physician was 6.5 (2.4) years, corresponding to the first 2 to 12 years in practice. Of the 3424 physicians, 21.5% had at least 1 complaint filed, and 17.1% had complaint(s) retained in their file after investigation. The majority (81.9%) of retained complaints were for attitude/communication and quality-of-care problems. Communication problems in management and inappropriate treatment/follow-up were the most common causes of quality-of-care complaints. Among the 696 retained complaints, none led to an immediate loss of license, 71 (10.2%) led to recommendations for additional counseling/training or discipline, and the remainder led to verbal and written warnings.

Lower CSE communication scores were associated with a higher rate of retained complaints, particularly in the lowest quartile of these scores (TABLE 4). The 853 physicians in the bottom communication score quartile had 236 retained complaints filed in their combined total of 5542 practice-years. This yielded an overall rate of 4.26 complaints per 100 practice-years compared with 2.51 per 100 practice-years for physicians in the top communication score quartile (Table 4). In multivariate models that adjusted for other physician characteristics, significantly higher complaint rates also were found for male vs female physicians, surgeons and primary care physicians vs medical subspecialists, and physicians practicing in Ontario vs those practicing in Quebec (Table 4). Even after adjustment for these characteristics, physicians in the lowest communication score quartile had an excess complaint rate of 1.75 per 100 practice-years compared with physicians in the top score quartile (adjusted relative risk [RR], 1.52; 95% CI, 1.30-1.78), and an excess complaint rate of 2.15 per 100 practice-years compared with the upper 3 quartiles (adjusted RR, 1.43; 95% CI, 1.22-1.68). The population attributable fraction indicated that 10.0% (95% CI, 6.0%-13.9%) of all retained

complaints were explained by physicians in the bottom communication score quartile.

There was no evidence of significant nonlinearity ( $P = .25$  for the GAM non-parametric test). According to the lin-

**Table 1.** Characteristics of the 3424 Physicians Taking the National Postgraduate Clinical Skills Examination Between 1993 and 1996 Who Were Licensed to Practice in Ontario and/or Quebec, Canada<sup>a</sup>

Physician Characteristics	No. (%)
Sex	
Female	1525 (44.5)
Male	1899 (55.5)
Age at the clinical skills examination, y	
<25	159 (4.6)
25-30	2451 (71.6)
>30	814 (23.8)
Mean (SD)	28.9 (4.7)
Undergraduate medical education	
Quebec/Ontario medical school	2655 (77.5)
Other Canadian medical school	349 (10.2)
International medical graduate	420 (12.3)
Postgraduate specialty program	
Family/general medicine	1393 (40.7)
Medical specialty	1481 (43.3)
Surgical specialty	550 (16.1)
Practice location	
Ontario	2263 (66.1)
Quebec	1009 (29.5)
Both provinces	152 (4.4)
Licensing examination performance <sup>b</sup>	Mean (SD) [Range]
Clinical skills examination	
Overall score	525.1 (79.9) [50-749]
Communication subscore	510.9 (91.1) [31-723]
Data acquisition subscore	508.8 (90.7) [19-875]
Problem-solving subscore	541.6 (98.5) [170-864]
Traditional written examination	
Overall score	526.5 (77.6) [338-787]
Multiple-choice questions subscore	524.4 (82.7) [278-793]
Clinical decision-making subscore	525.0 (75.8) [221-739]

<sup>a</sup>Percentages may not total 100% due to rounding.

<sup>b</sup>Scores standardized with a mean (SD) of 500 (100) for all first-time takers from Canadian medical schools in a given examination administration.

**Table 2.** Correlation Between Overall Scores and Subscores on the Medical Council of Canada Traditional Written and Clinical Skills Examinations<sup>a</sup>

Clinical Skills Examination	Traditional Written Examination <sup>b</sup>		
	Overall Score	Multiple-Choice Questions	Clinical Decision-Making Skills
Overall score	0.40	0.36	0.33
Communication	0.14	0.10	0.17
Data acquisition	0.23	0.21	0.16
Problem-solving	0.38	0.36	0.30

<sup>a</sup>For the traditional written examination scores, weighted Cronbach  $\alpha$  in different examination administrations was overall score, 0.92; multiple-choice questions, 0.91; and clinical decision making, 0.64. For clinical skills examination scores, weighted Cronbach  $\alpha$  was overall score, 0.77; communication, 0.41; data acquisition, 0.66; and problem solving, 0.54. Weights were based on the number of candidates taking the examination in each administration.<sup>36</sup>

<sup>b</sup>Disattenuated correlations ( $r$ )<sup>37</sup> between the scores for the traditional written examination and clinical skills examination were overall score, 0.47 and communication, 0.23; between the traditional written examination clinical decision-making score and the clinical skills examination score: overall score, 0.47 and communication, 0.43.

<sup>c</sup>Pearson product-moment correlation coefficients. All correlations were statistically significant ( $P < .001$ ).

ear model, a 2-SD decrease in communication score was associated with a relative 38% increase in the complaint rate (1.17 more complaints per 100 practice-years) (Table 4). The relationship between communication scores and complaint rate was significantly stronger in Quebec (RR, 1.84; 95% CI, 1.51-2.24) compared with Ontario (RR, 1.34;

**Table 3.** Frequency of Complaints by Type, Status, and Number of Physicians Among 3424 Physicians Followed Up for the First 2 to 12 Years of Practice and 22 585 Combined Practice-Years in Ontario and Quebec, Canada

Type of Complaint <sup>a</sup>	Proportion of Physicians With Complaints (N = 3424) <sup>b</sup>				Complaint Rate by Type (N = 22 585 Practice-Years)			
	≥1 Complaint		≥1 Retained Complaint		All Complaints		Retained Complaints	
	No.	% (95% CI)	No.	% (95% CI)	No.	Rate/100 Practice-Years (95% CI)	No.	Rate/100 Practice-Years (95% CI)
Attitude/communication	422	12.3 (11.1-13.4)	307	9.0 (8.0-10.0)	548	2.4 (2.2-2.6)	367	1.6 (1.4-1.8)
Communication problem in management of care <sup>c</sup>	356	10.4 (9.4-11.4)	239	7.0 (6.1-7.8)	357	1.6 (1.4-1.8)	240	1.1 (1.0-1.3)
Rude, abusive conduct to patients <sup>c</sup>	94	2.7 (2.2-3.2)	57	1.7 (1.3-2.1)	94	0.4 (0.3-0.5)	57	0.3 (0.2-0.4)
Quality of care	289	8.4 (7.5-9.3)	161	4.7 (4.0-5.4)	385	1.7 (1.5-1.9)	203	0.9 (0.8-1.0)
Inappropriate treatment/follow-up <sup>c</sup>	138	4.0 (3.3-4.7)	81	2.4 (1.9-2.9)	138	0.6 (0.5-0.7)	81	0.4 (0.3-0.5)
Inadequate assessment <sup>c</sup>	54	1.6 (1.2-2.0)	32	0.9 (0.5-1.2)	54	0.2 (0.2-0.3)	32	0.1 (0.1-0.2)
Professionalism	72	2.1 (1.6-2.6)	42	1.2 (0.8-1.6)	79	0.4 (0.3-0.5)	45	0.2 (0.2-0.3)
Office-related	37	1.1 (0.7-1.4)	28	0.8 (0.5-1.1)	39	0.2 (0.2-0.3)	30	0.1 (0.1-0.2)
Physician health problem	5	0.2 (0.1-0.3)	4	0.1 (0-0.2)	8	0.1 (0-0.1)	6	0.03 (0.01-0.07)
Other	50	1.5 (1.1-1.9)	42	1.2 (0.8-1.6)	57	0.3 (0.2-0.4)	45	0.2 (0.2-0.3)
Total	735	21.5 (20.1-22.9)	584	17.1 (15.8-18.4)	1116	4.9 (4.6-5.8)	696	3.1 (2.9-3.3)

Abbreviation: CI, confidence interval.

<sup>a</sup>Examples of professionalism included conflict of interest and advertising. Examples of office-related included inadequate records and office staff problems. Examples of physician health problems included mental health and alcohol-related behavior problems. Other complaints included those that were classified in multiple categories.

<sup>b</sup>118 Physicians had 2 or more retained complaints. The distribution of the 191 retained complaints against these physicians was similar to the distribution of all 696 retained complaints: communication, 86 (45%); quality of care, 68 (35.6%); professionalism, 24 (12.5%); office-related, 10 (5.2%); physician health problem, 1 (0.5%); and other, 2 (1.0%).

<sup>c</sup>The most prevalent subcategories of complaints within each category.

**Table 4.** Medical Council of Canada Clinical Skills Examination Communication Score and the Rate of Retained Complaints

	Population		Retained Complaint Rate		Relative Rate of Complaints Adjusted for Physician Characteristic <sup>a</sup>	
	No. of Physicians	Combined No. of Practice-Years	No.	Rate/100 Practice-Years (95% CI)	Relative Rate	P Value
					(95% CI)	
Communication score, by quartile <sup>b</sup>						
1	853	5542	236	4.26 (3.75-4.84)	1.52 (1.30-1.78)	<.001
2	847	5444	159	2.92 (2.50-3.41)	1.13 (0.96-1.32)	.29
3	867	5672	152	2.68 (2.29-3.14)	1.06 (0.90-1.24)	.63
4	857	5929	149	2.51 (2.14-2.95)	1 [Reference]	
By continuous score (per 2-SD decline in score)					1.38 (1.18-1.61)	<.001
Physician sex						
Female	1525	10 281	211	2.05 (1.79-2.35)	1 [Reference]	
Male	1899	12 305	485	3.94 (3.60-4.31)	1.64 (1.39-1.94)	<.001
Medical school						
Canadian	3004	19 615	580	2.96 (2.73-3.21)	1 [Reference]	
International	420	2970	116	3.91 (3.26-4.70)	1.11 (0.93-1.34)	.25
Specialty						
Medical specialty	1481	8162	163	2.00 (1.72-2.33)	1 [Reference]	
Family medicine or GP	1393	11 633	394	3.39 (3.07-3.74)	1.79 (1.49-2.16)	<.001
Surgical specialty	550	2790	139	4.98 (4.22-5.88)	2.43 (1.93-3.04)	<.001
Province of practice						
Ontario	2263	15 086	553	3.67 (3.38-3.99)	1 [Reference]	
Quebec	1009	6486	107	1.65 (1.37-1.99)	0.49 (0.40-0.61)	<.001
Both provinces	152	1014	36	3.55 (2.56-4.92)	1.00 (0.71-1.40)	.99

Abbreviations: CI, confidence interval; GP, general practice.

<sup>a</sup>Estimated by multivariate Poisson regression, adjusting for physician sex, specialty, country of training (Canada or international), and province.

<sup>b</sup>Cutoffs for the quartiles were first quartile, <457; second quartile, 457-518; third quartile, 519-575; fourth quartile, >575.

95% CI, 1.25-1.49). Physician sex and international medical graduate status were not significant modifiers of the communication score effect. Sensitivity analysis incorporating all complaints (retained and not retained) showed the same significant increase in the relative rate of complaints with declining communication score (6.55 per 100 practice-years in the lowest quartile compared with 4.78, 4.46, and 4.05 in the third, second, and upper quartile, respectively); however, the risk was smaller for all complaints (RR, 1.30; 95% CI, 1.22-1.39).

Among the CSE scores, only the communication score was significantly associated with complaint rates (TABLE 5). The CSE data acquisition and problem-solving scores had no relationship to complaint rate, including quality-of-care complaints. The CSE communication score was most strongly associated with the risk of communication complaints. The traditional written examination also was significantly associated with complaint rate, with the strongest association being for the clinical decision-making (CDM) score. The association between multiple-choice scores and complaint

rate was significant for overall retained complaints but not significant for communication or quality-of-care complaints. Statistically significant nonlinearity was found in the relationship between CDM scores and overall complaint rate ( $P=.02$ , for 3 *df* GAM test). The complaint rate increased with declining CDM scores between 600 and 450, with no systematic effect beyond this score range.

The CSE communication score, when added to a model that included traditional written examination CDM score, significantly improved the prediction of overall retained complaints and communication complaints, but not complaints about quality of care (Table 5). After adjustment for the traditional written examination CDM score, an additional 9.2% (95% CI, 4.7%-13.1%) of retained complaints and 11.2% (95% CI, 5.8%-16.9%) of communication complaints were explained by physicians in the bottom communication score quartile.

## COMMENT

In a longitudinal study of physicians who took the MCC clinical skills examination and entered practice in On-

tario and/or Quebec, scores obtained in patient-physician communication were statistically significant predictors of future complaints to medical regulatory authorities. The credibility of the association was strengthened by evidence of a linear relationship between complaint rates and communication scores, a slightly stronger association when the outcome was limited to communication complaints, consistency of the direction and statistical significance of the association in Ontario and Quebec, and the persistence of the association after adjustment for physician sex, specialty, international medical graduation status, and time in practice.

We observed a complaint rate of 0.0491 per physician. This rate is within the range of US state medical boards, where the mean complaint rate for all licensed physicians (including those with no complaints) varied from 0.02 per physician in Wisconsin to 0.20 per physician in Alabama between 2001 and 2003.<sup>41</sup> Similar to others, we found that communication problems were the most common reason for complaints<sup>42</sup>: 49.1% of complaints in our study compared with 55% of com-

**Table 5.** Scores on the Medical Council of Canada Qualifying Examinations and the Rate of Retained Complaints: Overall and by Type of Complaint

Examination Scores	Relative Rate of Retained Complaints by Examination Score After Adjustment for Physician Characteristics <sup>a</sup>					
	Any Retained Complaint		Communication Complaint		Quality-of-Care Complaint	
	Relative Rate (95% CI)	P Value	Relative Rate (95% CI)	P Value	Relative Rate (95% CI)	P Value
Clinical skills examination						
Overall score	1.19 (1.00-1.42)	.05	1.28 (1.00-1.64)	.05	1.06 (0.76-1.48)	.74
Communication score	1.38 (1.18-1.62)	<.001	1.43 (1.15-1.77)	.001	1.38 (1.03-1.86)	.03
Data acquisition score	0.98 (0.83-1.16)	.85	0.97 (0.78-1.22)	.82	1.00 (0.74-1.35)	.92
Problem-solving score	1.02 (0.88-1.19)	.76	1.13 (0.92-1.41)	.25	1.01 (0.76-1.33)	.97
Traditional written examination						
Overall score	1.39 (1.14-1.70)	.001	1.34 (1.01-1.76)	.04	1.54 (1.06-2.22)	.02
Multiple-choice score	1.25 (1.03-1.50)	.02	1.22 (0.94-1.57)	.14	1.29 (0.92-1.80)	.14
Clinical decision-making score	1.51 (1.25-1.84)	<.001	1.47 (1.13-1.92)	.004	1.77 (1.25-2.56)	.002
Clinical skills examination communication score, adjusted for clinical decision-making score <sup>b</sup>	1.32 (1.13-1.71)	<.001	1.37 (1.10-1.71)	.005	1.30 (0.97-1.75)	.09

Abbreviation: CI, confidence interval.

<sup>a</sup>Estimated by multivariate Poisson regression, adjusting for physician sex, specialty, country of training (Canada, international), and province, using the examination score as a continuous variable. Results presented as the change in relative rate per 2-SD decrease in score. A separate model was used to estimate the association of each score with retained complaints, adjusting for all of the same physician characteristics.

<sup>b</sup>Estimated by multivariate Poisson regression. Model includes communication score, clinical decision-making subscore of traditional written examination, physician sex, specialty, country of training (Canada, international), and province. Improvement in the fit of the model with clinical decision-making score alone and communication plus clinical decision-making score was assessed by likelihood ratio test.

plaints to 1 US state medical board between 1989 and 2000<sup>43</sup> and 74.7% in an investigation of hospital complaints between 2001 and 2003.<sup>6</sup>

Our results provide some feedback for medical educators and licensing authorities. Our study supports the predictive validity of providing a standardized assessment of communication skills prior to entry into practice. Almost 1 in every 5 physicians had a retained complaint filed with the provincial medical regulatory authorities in the first 2 to 12 years of practice. The risk of complaints was significantly greater among physicians in the lowest quartile of communication scores. This result suggests that direct observation and assessment of patient communication skills may be useful in identifying trainees who are more likely to experience difficulties in practice. Assessment of communication could play a role at different stages in training—to select candidates for medical school admission<sup>44</sup> or to identify trainees who may benefit from more intensive communication skill training, as these skills can be improved with training.<sup>45</sup>

In addition, our results suggest that a minimum passing standard should be established for communication on the CSE, as has been done in the US Step 2 Clinical Skills Examination.<sup>21</sup> To do so, the number of cases in which communication is assessed would need to be increased from the 3 to 4 cases to approximately 10 to 14 to obtain a sufficiently reliable score to make pass-fail decisions.<sup>46</sup> The MCC has already increased the number of cases in which communication is assessed to meet this reliability threshold.

Complaints were mainly associated with 2 subscores—clinical decision making and communication. Clinical decision-making assessment was specifically designed to select problems and test aspects of the decision-making process where physicians were more likely to make errors that would have an effect on patient outcome.<sup>34</sup> This approach to the selection of test material may explain why this component of the examination was

predictive of complaints, while the data collection and problem-solving components of the CSE were not. The key features approach to clinical decision-making assessment was first instituted by the MCC in 1992, and to our knowledge this is the first evaluation of its ability to predict future practice outcomes.<sup>47</sup> It may be useful to increase the use of key feature problems in traditional written assessment, as this format appears to be more predictive of quality-of-care complaints than ordinary multiple-choice questions. Selecting case and test elements for the national CSE on the same basis as key feature written problems also may be beneficial. The discriminating ability of data acquisition and problem-solving assessment on the CSE may be improved by selecting aspects of data collection that are critical for a given clinical problem, and where physicians tend to make errors.

Our study had several limitations. The poor-to-moderate reliability of the communication score component of the examination likely led to an underestimation of the strength of the relationship between communication and complaints.<sup>48</sup> The use of practice-years as a denominator for estimating the rate of complaints would not take into account differences between physicians in the frequency of patient contact, the type of patients, and the procedures performed, all of which may be associated with the risk of complaints. However, it seems unlikely that physicians with lower scores in communication would systematically seek out work activities and patient populations that are more likely to generate complaints.<sup>13</sup> On the other hand, higher rates of complaints that we found for surgeons, family physicians, and male physicians, even after adjustment for lower scores in communication, may be related to higher practice volume or differences in work activities or practice populations. As higher complaint and malpractice claim rates also have been found for these physician subgroups in other studies,<sup>1,10</sup> a better understanding of the contributing

factors would be important. Finally, we did not have information on language of greatest proficiency for the physician or language in which the test was taken, and could not include these factors in the analyses.

In summary, we found that communication and clinical decision-making ability were important predictors of future complaints to regulatory authorities. Current examinations could be modified to test these attributes more efficiently and at earlier points in the training process. Future research should examine whether remediation of communication problems can reduce complaints, and whether other indicators of the quality of practice could be assessed by a clinical skills examination.

**Author Contributions:** Dr Tamblin had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Tamblin, Abrahamowicz, Dauphinee, Jacques, Klass, Winslade.

**Acquisition of data:** Tamblin, Dauphinee, Wenghofer, Jacques, Klass, Smee, Girard, Bartman.

**Analysis and interpretation of data:** Tamblin, Abrahamowicz, Dauphinee, Wenghofer, Jacques, Klass, Blackmore, Girard, Du Berger, Bartman, Hanley, Buckeridge.

**Drafting of the manuscript:** Tamblin, Abrahamowicz, Dauphinee, Smee, Blackmore, Girard, Bartman, Hanley. **Critical revision of the manuscript for important intellectual content:** Tamblin, Dauphinee, Wenghofer, Jacques, Klass, Winslade, Du Berger, Hanley, Buckeridge.

**Statistical analysis:** Tamblin, Abrahamowicz, Girard, Du Berger, Bartman, Hanley, Buckeridge.

**Obtained funding:** Tamblin, Dauphinee.

**Administrative, technical, or material support:** Tamblin, Dauphinee, Wenghofer, Jacques, Klass, Smee, Blackmore, Winslade.

**Study supervision:** Tamblin, Abrahamowicz, Dauphinee, Jacques, Klass.

**Financial Disclosures:** None reported.

**Funding/Support:** The Medical Council of Canada (MCC) and the Canadian Institutes of Health Research (CIHR) provided funding for the study. The CIHR provided operating funds for the study and fellowship support for Dr Winslade.

**Role of the Sponsors:** The CIHR had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The MCC, the College of Physicians and Surgeons of Ontario, and the College of Physicians of Quebec, as organizations, were not involved in the approval of this article. Coauthors from the Medical Council of Canada (Dr Blackmore and Mss Smee and Bartman), the College of Physicians and Surgeons in Ontario (Drs Klass and Wenghofer), and College of Physicians of Quebec (Dr Jacques) provided oversight for the retrieval, linkage, and anonymization of data from their respective institutions, as well as input on the design and conduct of the study; collection, management, and interpretation of the data; and preparation, review, and approval of the manuscript. Mss Girard and Du Berger performed the statistical analysis.

## REFERENCES

1. Levinson W, Roger D, Mullooly J, Dull V, Frankel R. Physician-patient communication: the relationship with malpractice claims among primary care physicians and surgeons. *JAMA*. 1997;277(7):553-559.
2. Franks P, Fiscella K, Shields CG, et al. Are patients' ratings of their physicians related to health outcomes? *Ann Fam Med*. 2005;3(3):229-234.
3. Hickson GB, Clayton EW, Githens PB, Sloan FA. Factors that prompted families to file medical malpractice claims following perinatal injuries. *JAMA*. 1992;267(10):1359-1363.
4. Beckman HB, Markakis KM, Suchman AL, Frankel RM. The doctor-patient relationship and malpractice: lessons from plaintiff depositions. *Arch Intern Med*. 1994;154(12):1365-1370.
5. Kohatsu ND, Gould D, Ross LK, Fox PJ. Characteristics associated with physician discipline: a case-control study. *Arch Intern Med*. 2004;164(6):653-658.
6. Stelfox HT, Gandhi TK, Orav EJ, Gustafson ML. The relation of patient satisfaction with complaints against physicians and malpractice lawsuits. *Am J Med*. 2005;118(10):1126-1133.
7. Baldwin LM, Larson EH, Hart LG, Greer T, Lloyd M, Rosenblatt RA. Characteristics of physicians with obstetric malpractice claims experience. *Obstet Gynecol*. 1991;78(6):1050-1054.
8. Bovbjerg RR, Petronis KR. The relationship between physicians' malpractice claims history and later claims: does the past predict the future? *JAMA*. 1994;272(18):1421-1426.
9. Sloan FA, Mergenhagen PM, Burfield WB, Bovbjerg RR, Hassan M. Medical malpractice experience of physicians: predictable or haphazard? *JAMA*. 1989;262(23):3291-3297.
10. Hickson GB, Federspiel CF, Pichert JW, Miller CS, Gauld-Jaeger J, Bost P. Patient complaints and malpractice risk. *JAMA*. 2002;287(22):2951-2957.
11. Hickson GB, Pichert JW, Federspiel CF, Clayton EW. Development of an Early Identification and Response Model of Malpractice Prevention. *Law and Contemporary Problems*. 1997;60(1):7-29. Medical Malpractice: External Influences and Controls, part 1.
12. Vincent C, Young M, Phillips A. Why do people sue doctors? a study of patients and relatives taking legal action. *Lancet*. 1994;343(8913):1609-1613.
13. Hickson GB, Clayton EW, Entman SS, et al. Obstetricians' prior malpractice experience and patients' satisfaction with care. *JAMA*. 1994;272(20):1583-1587.
14. Moore PJ, Adler NE, Robertson PA. Medical malpractice: the effect of doctor-patient relations on medical patient perceptions and malpractice intentions. *West J Med*. 2000;173(4):244-250.
15. Papadakis MA, Teherani A, Banach MA, et al. Disciplinary action by medical boards and prior behavior in medical school. *N Engl J Med*. 2005;353(25):2673-2682.
16. Fiscella K, Meldrum S, Franks P, et al. Patient trust: is it related to patient-centered behavior of primary care physicians? *Med Care*. 2004;42(11):1049-1055.
17. Stillman P, Swanson D, Regan MB, et al. Assessment of clinical skills of residents utilizing standardized patients. *Ann Intern Med*. 1991;114(5):393-401.
18. AAMC. Report III: Contemporary Issues in Medicine: Communication in Medicine. <http://www.aamc.org/meded/msop/>. Washington, DC: Association of American Medical Colleges; 2006:1-29. Medical School Objectives Project. Accessed August 10, 2007.
19. Reznick RK, Blackmore D, Dauphinee WD, Rothman AI, Smees S. Large-scale high-stakes testing with an OSCE: report from the Medical Council of Canada. *Acad Med*. 1996;71(1)(suppl):S19-S21.
20. Whelan G. Educational Commission for Foreign Medical Graduates: lessons learned in a high-stakes, high-volume medical performance examination. *Med Teach*. 2000;22(3):293-296.
21. Papadakis MA. The Step 2 clinical-skills examination. *N Engl J Med*. 2004;350(17):1703-1705.
22. USMLE. Harris Interactive Poll, Results from: United States Medical Licensing Examination. <http://www.usmle.org/news/Step2CSNews/Harris.asp>. Accessed August 10, 2007.
23. Hoppe RB. *Performance Assessment: Why Should Medical Schools Care?* Henry R, Mavis B, Mullan P, Pryor O, Solomon D, eds. East Lansing: Office of Medical Education Research and Development, Michigan State University; 2002.
24. Bogdonoff MD. The clinical-skills examination. *N Engl J Med*. 2004;351(5):507-509.
25. Musco S. The clinical-skills examination. *N Engl J Med*. 2004;351(5):507-509.
26. Henry SG. Playing doctor. *JAMA*. 2005;294(17):2138-2140.
27. Hallock JA, Melnick DE, Thompson JN. The step 2 clinical skills examination. *JAMA*. 2006;295(10):1123-1124.
28. Federation of State Medical Boards. Elements of a modern state medical board. [http://www.fsmb.org/pdf/GRPOL\\_Elements\\_Modern\\_Medical\\_Board.pdf](http://www.fsmb.org/pdf/GRPOL_Elements_Modern_Medical_Board.pdf). 2006. Accessed August 10, 2007.
29. Federation of State Medical Boards. A guide to the essentials of a modern medical practice act—tenth edition. [http://www.fsmb.org/pdf/2003\\_grpol\\_modern\\_medical\\_practice\\_act.pdf](http://www.fsmb.org/pdf/2003_grpol_modern_medical_practice_act.pdf). 2003. Accessed August 10, 2007.
30. Dauphinee WD. Self regulation must be made to work. *BMJ*. 2005;330(7504):1385-1387.
31. Federation of Medical Regulatory Authorities of Canada. <http://www.fmrac.ca>. 2007. Accessibility verified July 16, 2007.
32. Federation of State Medical Boards. <http://www.fsmb.org>. 2007. Accessibility verified July 16, 2007.
33. Mandin H, Dauphinee WD. Conceptual guidelines for developing and maintaining curriculum and examination objectives: the experience of the Medical Council of Canada. *Acad Med*. 2000;75(10):1031-1037.
34. Page G, Bordage G, Allen T. Developing key-feature problems and examinations to assess clinical decision-making skills. *Acad Med*. 1995;70(3):194-201.
35. Scherpier AJA, van der Vleuten C, Rethans J-J, van der Steeg AFW, eds. *Optimizing the Input of Physician Examiners in Setting Standards for a Large Scale OSCE: Experience With Part II of the Qualifying Examination of the Medical Council of Canada*. Dordrecht, the Netherlands: Kluwer Academic Publishers; 1997.
36. Osborne JW. Effect sizes and disattenuation of correlation and regression coefficients: lessons. <http://pareonline.net/getvn.asp?v=8&n=118>[11]. 2003. Practical Assessment, Research & Evaluation. Accessed August 10, 2007.
37. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. London, England: Chapman & Hall; 1990.
38. Hanley JA. A heuristic approach to the formulas for population attributable fraction. *J Epidemiol Community Health*. 2001;55(7):508-514.
39. Signorini D. Sample size for Poisson regression. *Biometrika*. 1991;78(2):446-450.
40. Edelstein RA, Reid HM, Usatine R, Wilkes MS. A comparative study of measures to evaluate medical students' performance. *Acad Med*. 2000;75(8):825-833.
41. Roberts RG, Friedsam D, Beasley JW, Helstad C, Moberg DP. The state of quality reports: comparing states by their rankings. *WJM*. 2006;105(8):60-66.
42. Federation of State Medical Boards. State medical board disciplinary reports: 2005; trends in physician regulations—2006. [http://www.fsmb.org/pdf/PUB\\_FSMB\\_Trends\\_in\\_Physician\\_Regulation\\_2006.pdf](http://www.fsmb.org/pdf/PUB_FSMB_Trends_in_Physician_Regulation_2006.pdf). April 2006. Accessibility verified July 16, 2007.
43. Halperin EC. Grievances against physicians: 11 years' experience of a medical society grievance committee. *West J Med*. 2000;173(4):235-238.
44. Lumsden MA, Bore M, Millar K, Jack R, Powis D. Assessment of personal qualities in relation to admission to medical school. *Med Educ*. 2005;39(3):258-265.
45. Yedidia MJ, Gillespie CC, Kachur E, et al. Effect of communications training on medical student performance. *JAMA*. 2003;290(9):1157-1165.
46. Ben-David MF, Klass DJ, Boulet J, et al. The performance of foreign medical graduates on the National Board of Medical Examiners (NBME) standardized patient examination prototype: a collaborative study of the NBME and the Educational Commission for Foreign Medical Graduates (ECFMG). *Med Educ*. 1999;33(6):439-446.
47. Page G, Bordage G. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Acad Med*. 1995;70(2):104-110.
48. Cleary TA, Linn RL, Walster GW. Effect of reliability and validity on power of statistical tests. *Sociol Methodol*. 1970;2:130-138.