Association Between Licensure Examination Scores and Practice in Primary Care

Robyn Tamblyn, PhD
Michal Abrahamowicz, PhD
W. Dale Dauphinee, MD
James A. Hanley, PhD
John Norcini, PhD
Nadyne Girard, MSc
Paul Grand'Maison, MD
Carlos Brailovsky, MD, PhD

Context Standards for licensure are designed to provide assurance to the public of a physician's competence to practice. However, there has been little assessment of the relationship between examination scores and subsequent practice performance.

Objective To determine if there is a sustained relationship between certification examination scores and practice performance and if licensing examinations taken at the end of medical school are predictive of future practice in primary care.

Design, Setting, and Participants A total of 912 family physicians, who passed the Québec family medicine certification examination (QLEX) between 1990 and 1993 and entered practice. Linked databases were used to assess physicians' practice performance for 3.4 million patients in the universal health care system in Québec, Canada. Patients were seen during the follow-up period for the first 4 years (1993 cohort of physicians) to 7 years (1990 cohort of physicians) of practice from July 1 of the certification examination to December 31, 1996.

Main Outcome Measures Mammography screening rate, continuity of care index, disease-specific and symptom-relief prescribing rate, contraindicated prescribing rate, and consultation rate.

Results Physicians achieving higher scores on both examinations had higher rates (rate increase per SD increase in score per 1000 persons per year) of mammography screening (β for QLEX, 16.8 [95% confidence interval {CI}, 8.7-24.9]; β for Medical Council of Canada Qualifying Examination [MCCQE], 17.4 [95% CI, 10.6-24.1]) and consultation (β for QLEX, 4.9 [95% CI, 2.1-7.8]; β for MCCQE, 2.9 [95% CI, 0.4-5.4]). Higher subscores in diagnosis were predictive of higher rates in the difference between disease-specific and symptom-relief prescribing (β for QLEX, 3.9 [95% CI, 0.9-7.0]; β for MCCQE, 3.8 [95% CI, 0.3-7.3]). Higher scores of drug knowledge were predictive of a lower rate (relative risk per SD increase in score) of contraindicated prescribing for MCCQE (relative risk, 0.88; 95% CI, 0.77-1.00). Relationships between examination scores and practice performance were sustained through the first 4 to 7 years in practice.

Conclusion Scores achieved on certification examinations and licensure examinations taken at the end of medical school show a sustained relationship, over 4 to 7 years, with indices of preventive care and acute and chronic disease management in primary care practice.

JAMA. 2002;288:3019-3026 www.jama.com

THE MEDICAL PROFESSION ensures the basic competence of physicians by requiring them to pass licensing and certifying examinations.[1] Although it is generally assumed that these examinations predict how physicians will practice in the future,[2] the data in support of this assumption are mostly indirect. Physicians who have more training in a discipline are more knowledgeable[3,4] and achieve higher scores in their respective discipline on recertification examinations.[5] More knowledgeable physicians are more likely to adhere to evidence-based guidelines in the delivery of care[6,7] and achieve better patient outcomes.[6] Certification status, which represents pass/fail status on certification examinations, is an important predictor of quality of care.[8,9]

It is unknown, however, whether scores achieved by physicians with the same training and specialty are predictive of future performance. A prior study[8] found that scores on an internal medicine certification examination predicted colleagues' ratings of the quality of care delivered by internists 5 to 8 years later. However, little is known about the relationship between examination scores and more objective measures of quality of care. Assessing this relationship is relevant because important gaps exist between optimal and actual practice in the delivery of preventive care,[10] in the management of acute and chronic disease,[7,11,12] and in the quality of drug treatment.[13] All of these are deficiencies that could potentially be predicted by licensing examinations.

We previously reported that physicians who achieved higher scores on the Québec family medicine certification examination were more likely to refer women for mammography screening, to prescribe more disease-specific medication and fewer symptomatic and contra-

Author Affiliations: Departments of Medicine (Dr Tamblyn) and Epidemiology and Biostatistics (Drs Tamblyn, Abrahamowicz, and Hanley, and Ms Girard), McGill University, Montreal, Québec; Medical Council of Canada, Ottawa, Ontario (Dr Dauphinee); Foundation for Advancement of International Medical Education and Research, Philadelphia, Pa (Dr Norcini); Department of Family Medicine, University of Sherbrooke, Sherbrooke, Québec (Dr Grand'Maison); and Centre d'évaluation des sciences de la santé, University of Laval, Ste-Foy, Québec (Dr Brailovsky).
Corresponding Author and Reprints: Robyn Tamblyn, PhD, McGill University, Morrice House, 1140 Pine Ave W, Montréal, Québec, Canada H3A 1A3 (e-mail: robyn.tamblyn@mcgill.ca).

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

indicated drugs, and to refer more of their patients for consultation.[14] However, assessment of these outcomes was limited to the first 18 months of practice. We used this opportunity to determine if the association between family medicine certification examination scores and practice performance persisted with increasing practice experience.[14] We could not find any studies that explored the relationship between earlier licensing examinations taken at the end of medical school and future practice performance. We investigated whether such scores predict clinical behaviors 4 to 7 years later.

METHODS

Context

All Canadian provinces provide a universal health insurance program that covers the costs of medical care for provincial residents. In Québec, 14 500 physicians provide services to 7.4 million residents of the province, for whom 92% of services and 93% of physicians are paid by the Québec health insurance agency (Régie de l'assurance maladie du Québec; RAMQ) on a fee-for-service basis.[15]

Design and Study Population

A cohort of all family physicians who passed the Québec family medicine certification examination between 1990 and 1993, and entered fee-for-service practice in Québec, was followed up for the first 4 to 7 years of practice. Annual measures of each physician's practice performance were used to test associations between examination scores and practice performance. Salaried physicians were excluded because there was no accurate way to identify all their patients, as were physicians who subsequently trained in another specialty. Potentially eligible physicians were identified by the Québec College of Physicians, and physicians' license numbers were provided to the RAMQ to retrieve data on practice activity. Annual measures of each physician's clinical behaviors were used to test associations between examination scores and practice performance. Data on all patients seen by each physician, for each year of practice, were retrieved from the RAMQ to assess annual practice activity and case-mix differences between physician practices. For each cohort, the follow-up period was between July 1 of the certification examination and December 31, 1996, providing practice assessment for 4 years (1993 cohort) to 7 years (1990 cohort). The study protocol was approved by the institutional review board at McGill University, the Provincial Access to Information Office, and the RAMQ legal counsel.

Family Medicine Certification Examination

The Québec Licensing Examination (QLEX) comprises the College of Family Physicians of Canada Certification Examination (CFPCex) and the Québec Objective Structured Clinical Examination (OSCE).[16] Between 1990 and 1993, the CFPCex assessed diagnosis, management, and prevention with multiple-choice questions and short-answer management problems, and assessed communication skills with a simulated patient. The OSCE measured clinical problem-solving skills by direct observation of performance in 25 standardized patient encounters, rated by physician examiners.[16] To pass the examination, a score of 60% was required in diagnosis, management, and communication; 50% in prevention; and an OSCE score greater than 2 SDs below the mean. The overall reliability of examination scores varied from a low of 0.27 for prevention to a high of 0.72 for the OSCE.[14] Scores were standardized to adjust for differences in the difficulty of examinations using the reference group approach.[14] The College of Physicians provided the license number, medical school, and encrypted examination scores to the RAMQ, who then retrieved data on practice activity and linked this with score files through encrypted physician and patient identifiers.

Medical Council of Canada Qualifying Examination

The Medical Council of Canada Qualifying Examination (MCCQE) is used to test an individual's competence to enter postgraduate training. It is generally taken during the final year of medical school. A passing score on the MCCQE is required for unrestricted licensure in all Canadian provinces except Québec, although most Québec graduates take the examination. The 1988-1991 examinations consisted of 450 multiple-choice questions, which tested knowledge in medicine, surgery, obstetrics/gynecology, psychiatry, pediatrics, and preventive medicine, and also included 28 to 35 patient management problems that tested competence in clinical problem solving.[17] To enable comparisons between the MCCQE and the Québec certification examination, the 1258 multiple-choice questions used between 1988 and 1991 were reclassified by the test committees as diagnosis, management, or other knowledge. Test committees also identified management items that tested drug-specific knowledge. Standardized ability scores were created for each subscore using the same reference group approach as the certification examination.[18] The overall score reliability was 0.92, and subscore reliabilities were 0.71 for prevention, 0.63 for diagnosis, 0.73 for management, and 0.48 for drug knowledge. The Medical Council of Canada linked score data to the Québec cohort file by name, sex, and birth date.

Practice Assessment

Data Sources and Retrieval. Four previously validated[19] health administrative databases, linked by encrypted beneficiary identifiers, were used to assess practice performance.[20,21] The registrant database provided patients' age, sex, postal code, and date of death. The medical services database provided type, location, diagnosis, treating and referring physician, and date of all services delivered on a fee-for-service basis. The prescription database provided drug, dose, duration, prescribing physician, and date of each prescription dispensed for patients aged 65 years or older. The hospitalization database provided records of all hospital discharges including discharge diagnoses and admission and discharge dates. The 1991 census data was linked by 6-digit postal code to the reg-

 ©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

istrant database to measure mean family income and educational achievement in the residential area of each patient.[22] For each physician, the medical services claims files were used to identify all patients seen by the physician from the licensure date to December 31, 1996. For each patient, the RAMQ retrieved demographic data, as well as all medical services, prescriptions, and hospitalizations provided during the follow-up period. We also obtained data for the year prior to the first contact with the study physician to provide an independent assessment of patient characteristics that could not have been influenced by the study physician.[14]

Indicators of Practice Performance. We assessed 6 performance indicators in each follow-up year, selected on the basis of unexplained practice variation, and/or their association with the outcomes or costs of care.[14] First, a mammography screening rate was used to assess preventive care, because physicians who screen for breast cancer are also more likely to perform other preventive services.[23] Second, continuity of care was selected because of its importance in prevention and chronic disease management,[24-27] and to test the hypothesis that communication and management skills are predictive of better continuity.[28,29] Third, the differences between disease-specific and symptom-relief prescribing rate and contraindicated prescribing rate were used as indicators of the quality of acute and chronic disease management because variation in disease-specific relative to symptom-relief prescribing rate[30] has been linked to diagnostic competence.[14] Fourth, contraindicated prescribing, which accounts for 20% of drug-related adverse events, may be caused by deficiencies in physician knowledge.[31] Finally, consultation rate was used as an indicator of resource use because referral determines access to higher cost specialty care,[32] and primary care physicians with higher self-reported competence appear more likely to refer patients for specialty consultation.[33,34]

Annual mammography screening rate was defined as the proportion of eligible women in the primary care practice who were referred for a bilateral mammogram by the physician. Eligible women were between age 50 and 69 years, had no prior diagnosis of breast cancer, breast disease, or diagnostic mammogram, were due for screening, and were not receiving primary care from a gynecologist or obstetrician.

Annual continuity of care was defined as the mean proportion of visits that were made to or referred by the study physician by all patients in the primary care practice population. Using the medical services claims, we determined the proportion of all visits in the year for each patient that were made either to the study physician or to other physicians based on study physician referral. Each patient's proportion was weighted by the square root of the total number of visits made by the patient in the year. An overall mean annual continuity of care index for each physician's practice was calculated as the weighted mean of individual patient's proportions.

Mammography screening and continuity of care were assessed only in the primary care practice population, which consisted of patients seen in an office or clinic in the respective year, for whom the study physician had provided an annual physical or major (3 system) assessment. Consultation was assessed in all ambulatory patients for whom the study physician had billed at least 1 outpatient, office practice, or emergency department visit. Prescribing rates were assessed in ambulatory patients aged 65 years or older for whom complete information was available on all prescriptions dispensed.

Annual difference between disease-specific and symptom-relief prescribing rates was determined by examining medications prescribed by the study physician to all elderly patients. *Disease-specific medication* was defined as drugs that would rarely be prescribed without an investigation-confirmed disease (eg, anticoagulants, anticonvulsants, antidepressants, antihypertensives, medications for cardiovascular disease, asthma, and Parkinson disease, corticosteroids, diuretics, antiglaucoma medication, hypoglycemic medication, and thyroid medication). *Symptom-relief medication* was defined as drugs that relieve symptoms, but have little impact on the disease process (eg, nonsteroidal anti-inflammatory medications, benzodiazepines, low-dose narcotic analgesics) using the McGavock classification.[30]

Annual contraindicated prescribing rate was the proportion of ambulatory elderly patients for whom the study physician prescribed a relatively contraindicated medication. These were defined by an updated expert review[35,36] as 30 drugs that should be avoided in elderly patients because of possible toxic effects.

Annual consultation rate was the proportion of all ambulatory patients in the respective calendar year referred, at least once, to a specialist by the respective primary care physician, based on medical services claims. To be reimbursed for a consultation visit, the consultant must record a valid license number for the referring physician.

Time in Practice. Physicians' preceding practice experience was represented as a time-dependent covariate, based on a count of the number of previous months that the physician had at least 1 fee-for-service billing in the medical services claims file.

Case-Mix Assessment

Relevant data were retrieved for each eligible patient for the 12-month period preceding the first contact with the study physician. Individual characteristics of all patients seen in a given year by the physician were then aggregated to adjust for between-physician differences in practice population case-mix. Each practice population was characterized by age and sex distribution, mean family income, educational achievement, geographic access to health care,[14] propensity for the use of health care services,[37] comorbidity (Charlson index),[38] and hospitalization rates in the previous year. These annual case-mix measures were modeled as time-dependent covariates. For contraindicated prescribing, no adjustments were made for case-mix because the prescription of relatively contraindicated medication is rarely justified by patient characteristics.[14]

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

Statistical Analysis

Relationships between examination scores and practice performance were tested using multiple linear and Poisson regression for repeated measures using generalized estimating equations.[39] Physicians were the unit of analysis. Outcome variables were the 5 annual measures of practice performance. An autoregressive first-order correlation structure of residuals was used to characterize the interdependence between annual performance measures for each physician. To reduce imprecision in outcome measurement, physicians were excluded in years in which they had fewer than 5 patients, and the logarithm of the number of patients in the practice population in a given year was used as a weight in the analysis. Linearity assumptions were evaluated by testing the statistical significance of the quadratic component. Poisson regression was used to assess rates of contraindicated prescribing because such events were rare. The SEs were empirically estimated to account for overdispersion. Annual measures of case-mix were included in all analyses, as were indicators of the medical school to conservatively assess the impact of variation in examination scores within each medical school. To determine whether the association between examination scores and practice outcomes diminished over time, the interactions between examination score and practice experience were tested. For significant relationships between examination score and outcomes that persisted over time, the cumulative difference in the number of outcomes per 1000 patients followed up over the first 5 years in practice by high-scoring (2 SDs above the mean) rather than low-scoring (2 SDs below the mean) physicians was estimated. Regression coefficients for the score-outcome relationship were used to determine the expected annual difference in rates corresponding to a 4-SD difference in examination scores, and then the result was multiplied by 5 to estimate the cumulative impact over 5 years. $P<.05$ was used as the level of statistical significance. We used SAS statistical software to perform our analyses (Version 8.0, SAS Institute Inc, Cary, NC).

Table 1. Physician Characteristics and Examination Scores for 912 Québec Family Physicians*

	No. (%)
Sex	
Male	382 (41.9)
Female	530 (58.1)
Medical school	
Québec	
A	312 (34.2)
B	260 (28.5)
C	69 (7.6)
D	164 (18.0)
Other part of Canada or in United States	56 (6.1)
Other international	51 (5.6)
Postgraduate training same as undergraduate medical school	628 (81)
Certification year	
1990	199 (21.8)
1991	254 (27.9)
1992	271 (29.7)
1993	188 (20.6)
Took MCCQE examination	782 (85.8)
	Mean (SD) [Range]
QLEX score	
Overall	−0.03 (1.05) [−3.97 to 2.76]
Subscore	
Diagnosis	0.01 (1.07) [−3.97 to 4.40]
Management	−0.04 (1.03) [−4.53 to 2.47]
Prevention	−0.02 (1.06) [−4.63 to 3.04]
Clinical assessment	−0.01 (0.96) [−2.81 to 2.91]
MCCQE score	
Overall	−0.22 (0.89) [−4.31 to 2.24]
Subscore	
Diagnosis	0.02 (0.92) [−3.64 to 2.55]
Management	0.21 (0.89) [−3.10 to 2.67]
Prevention	−0.19 (0.86) [−3.66 to 1.95]
Drug knowledge	−0.04 (0.92) [−3.04 to 2.93]
Clinical assessment	−0.83 (1.20) [−5.97 to 2.03]

*Examination scores are standardized to a mean of zero representing the average score for first-time takers of the examination from North American medical schools. MCCQE indicates Medical Council of Canada Qualifying Examination; QLEX, Québec Certification Examination.

RESULTS

Certification Examination

Between 1990 and 1993, a total of 944 family physicians passed the QLEX, 920 (97.5%) started practice in Québec, 912 (96.7%) entered a fee-for-service practice, of whom 58.1% were female (TABLE 1). Overall, 85.8% of physicians took the MCCQE during the final year of medical school. Québec medical school graduates had modestly lower MCCQE scores in comparison with other Canadian graduates (mean [SD], −0.04 [0.84] vs 0.09 [0.92]; $P<.001$). Graduates who entered family medicine or general practice training in Canada had slightly lower standardized scores than those entering specialty programs (mean [SD], 0.02 [0.95] vs 0.13 [0.84]; $P<.001$). Québec family physicians who did not take the MCCQE had lower certification examination scores than those who did (mean [SD], −0.76 [1.18] vs −0.03 [1.02]; $P<.001$). Mean scores achieved on both the QLEX and MCCQE were equivalent or slightly lower than in the reference group of first-time takers, with a typical range of 6 to 7 SDs (Table 1). The Pearson correlation between the MCCQE and QLEX was 0.55, and ranged from 0.26 (clinical problem solving) to 0.49 (management).

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

Physicians practiced in a mean of 2.6 (year 1-2) to 3.2 (year 5-7) different types of settings (TABLE 2). During the first 4 to 7 years of practice, an increasing proportion of physicians moved their primary practice base from rural to urban populations, and established private office practice. After the first 2 years, practice size, number of days worked per year, and visits per day continued to increase, but at a slower rate.

The 912 physicians billed for 3.4 million different patients (45.9% of the Québec population), of whom 1.4 million were in their primary care practice populations, 385 321 were elderly, and 119 866 were women eligible for mammography screening.

The relationship between the QLEX and mammography screening was sustained over the first 4 to 7 years in practice (TABLE 3). The significant interaction between certification examination score and practice experience indicated that the strength of the relationship increased over time. For every SD increase in QLEX score, the mammography screening rate increased by 13.3 women per 1000 in years 1 to 2 of practice, 21.7 by years 3 to 4, and 19.2 by years 5 to 7. The persistence of this relationship means that, during the first 5 years of practice, high-scoring physicians would be expected to order 347 more mammograms per 1000 women than low-scoring physicians. In contrast, there was no relationship between QLEX scores and continuity of care. Overall, however, family physicians provided or coordinated only 28.3% of all visits made by patients in their primary care practice population.

Consultation rate showed a persistent, but modest, association with QLEX score over the first 7 years of practice (Table 3). Each SD increase in score was associated with an additional 2.92 referrals for specialty consultation per 1000 patients seen per year, resulting, over the first 5 years, in 58 more referrals per 1000 patients by high-scoring than low-scoring physicians.

Diagnosis and management subscores of the QLEX were the only significant predictors of prescribing outcomes. Higher scores were associated with higher rates of disease-specific relative to symptom-relief prescribing, and a lower risk of contraindicated prescribing, although the latter did not achieve statistical significance (Table 3). The strength and significance of the associations between examination scores and practice outcomes increased when medical school was excluded from the model, because there were systematic differences in scores for graduates from different medical schools.

Qualifying Examination

Scores on the MCCQE, taken at the end of medical school, had a similar pattern of relationships as the certification examination (TABLE 4). For example, the increase in mammography screening rate per SD increase in score was 16.8 per 1000 for the MCCQE score (Table 4) and 17.4 per 1000 for the overall QLEX score (Table 3). The diagnosis subscore was the strongest predictor of differences in the rates of disease-specific and symptom-relief prescribing in both the MCCQE and the QLEX. The drug knowledge subscore was the only significant predictor of contraindicated prescribing—reducing the risk of contraindicated prescribing by 12% per SD increase in score. During the first 5 years, a high-scoring physician would be expected to write 85 fewer contraindicated prescriptions per 1000 elderly patients than a low-scoring physician.

COMMENT

We examined the relationship between licensure and certification examination scores and practice performance in a sample of newly certified family physicians. Linked databases

Table 2. Practice Setting and Workload Characteristics for Family Physicians in the First 4 to 7 Years of Practice

	Physician Practice Year		
	1-2 (n = 912)	3-4 (n = 841)	5-7 (n = 516)
	No. (%)		
Practice setting*			
Hospital-based			
Outpatient clinic	542 (59)	439 (52)	255 (49)
Emergency department	629 (69)	582 (69)	361 (70)
Inpatient unit	611 (67)	565 (67)	361 (70)
Intensive care unit	289 (32)	265 (31)	154 (30)
Long-term care	255 (28)	199 (24)	103 (20)
Private office practice	629 (69)	611 (73)	430 (83)
Community clinic	104 (11)	88 (10)	39 (8)
	Mean (SD)		
Patient geographic distribution, %†			
Urban	41.2 (43.7)	43.2 (44.1)	46.5 (44.2)
Intermediate	34.1 (42.5)	34.0 (42.5)	34.0 (42.3)
Rural-remote	24.7 (40.0)	22.8 (39.2)	19.5 (36.9)
Practice settings per year	2.6 (2.0)	3.0 (1.9)	3.2 (1.7)
Workload			
Patients seen per year	1174 (1259)	1861 (1470)	2147 (1385)
Work days per year	97 (92)	150 (96)	174 (79)
Visits per work day	16 (7)	18 (7)	18 (7)

*Most physicians practiced in multiple practice settings (eg, in year 1-2, the average number of different practice settings was 2.6 per physician). For this reason, the number (percentage) of physicians practicing in each type of practice setting adds to more than 100%, as most physicians were represented in more than 1 category.

†The residence of each patient in a physician's practice population was categorized as urban (resided in the regions of Montréal, Québec, Laval, or Montérégie), intermediate (resided in the regions of Lanaudières, Estrie Saguenay-Lac-St-Jean, Laurentides, Mauricie-Bois-Francs, or Outaouais), or rural-remote (resided in the regions of Chaudières-Appalaches, Abitibi-Témiscamingue, Gaspésie, Bas-Saint-Laurent, Côte-Nord, Nord-du-Québec, Kativik Terres-cries-de la Baie-James). For each physician, the proportion of patients from urban, intermediate, and rural-remote regions was determined for the first 1 to 2, 3 to 4, and 5 to 7 years of practice. The mean represents the average proportion of patients in the practices of physicians in the cohort who resided in urban, intermediate, and rural-remote locations.

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

from a universal health care system allowed us to follow-up all licensed physicians, and all their patients in their first 4 to 7 years of practice. Thus, we minimized selection biases that could confound the assessment of relationships. We were also able to assess the population impact of potential licensure regulations. The most important findings of this study were that the relationships between certification examination scores were sustained through the first 4 to 7 years of practice, and that examinations taken in the final year of medical school were also significant predictors of practice performance. However, the use of administrative databases had limitations. The databases restricted the type of performance indicators that could be measured and we were limited in our ability to adjust for confounding by case-mix, but as case-mix was not associated with examination scores, it is unlikely that this introduced substantial bias in the results. These results have several implications for licensing bodies.

Such outcome data could be incorporated into the establishment of examination passing standards so that standard-setting participants can weigh their usual judgments of test content and pass rates against the consequences for patients.[40-42] For example, an increase in the passing criterion of only 1 SD in drug knowledge would

Table 3. Association Between Family Medicine Certification Examination Scores and Practice Performance in the First 4 to 7 Years of Practice

			Change in Outcome per SD Increase in Score					
			All Practice Years		Practice Year†			
Outcome	Mean (SD) Rate per 1000 Patients	Type of Certification Examination Score	β (95% CI)*	*P* Value	1-2	3-4	5-7	*P* Value for Interaction‡
Preventive care								
Mammography screening rate among eligible women	117.4 (136.6)	Overall	17.37 (10.6 to 24.1)	<.001	13.3	21.7	19.2	.005
		Prevention	15.61 (8.9 to 22.3)	<.001	17.9	15.9	12.7	.03
		Clinical assessment	14.22 (6.9 to 22.6)	.001	11.7	18.4	13.4	.11
Coordination of care								
Continuity of care, % of visits	28.3 (11.0)	Overall	0.2 (–0.3 to 0.6)	.46	0.2	0.03	0.05	.27
		Communication	0.05 (–0.4 to 0.5)	.85	0.1	–0.1	0.1	.64
		Management	0.2 (–0.3 to 0.6)	.44	0.4	0.09	–0.2	.04
Resource use								
Consultation rate	76.6 (52.3)	Overall	2.92 (0.4 to 5.4)	.02	2.7	2.4	3.3	.74
Acute and chronic disease management								
Symptom-relief prescription rate among elderly patients	148.9 (116.7)	Overall	–2.83 (–8.7 to 3.0)	.34	–2.7	–2.8	–2.9	.66
		Diagnosis	–2.56 (–7.9 to 2.8)	.34	–2.0	–1.8	–3.9	.35
		Management	–5.97 (–12.0 to 0)	.05	–5.8	–6.2	–5.9	.32
Disease-specific prescription rate minus symptom-relief prescription rate	–1.45 (76.2)	Overall	2.89 (–0.6 to 6.4)	.10	1.5	2.9	5.3	.24
		Diagnosis	3.94 (0.9 to 7.0)	.01	3.0	3.6	6.1	.15
		Management	2.42 (–0.9 to 5.7)	.15	1.6	3.3	2.7	.95
Contraindicated prescriptions among elderly patients§	32.6 (43.3)	Overall	0.97 (0.9 to 1.1)	.62	0.97	0.96	0.99	.52
		Diagnosis	0.99 (0.9 to 1.1)	.97	0.98	0.99	1.04	.16
		Management	0.94 (0.8 to 1.0)	.26	0.94	0.93	0.95	.27

*The regression coefficient β represents the estimated change in the rate or value of the practice outcome per SD increase in score in the first 4 to 7 years in practice with a 95% confidence interval (CI). In these overall models, the interaction term to test the potential modification of the magnitude of the effect between practice outcome and examination scores in relationship with the number of months in practice is not included. In instances in which there was a significant interaction between examination score and months in practice (ie, mammography screening rate), the estimates for each category of years in practice provide a more appropriate estimate of the effects. Each β was estimated by a multivariate regression model within a generalized estimating equation framework, in which physician was the unit of analysis and annual assessment of outcome rates/values, were represented as repeated measurements for each physician. Observations were weighted by the logarithm of each physician's annual practice size. The estimate of the examination score, practice outcome relationship was adjusted for differences in annual practice case-mix including age and sex structure, socioeconomic status, geographic access to health care, comorbidity, and propensity to use health care services based on data for individual practice patients in the year prior to outcome assessment. The only exception was for contraindicated prescribing in which practice size was used to weight estimate regression coefficients and medical school was included, but practice case-mix covariates were not included as these attributes of the practice population would rarely justify contraindicated prescribing in the elderly. When medical school was excluded from the regression models, the magnitude of the association between examination scores and practice outcomes increased because some medical schools had systematically lower scores than others. If the analysis were based on usual practice, in which pass-fail decisions are made irrespective of medical school, the overall predictive relationship between examination scores and outcomes would have been higher: mammography screening and overall score (β, 19.3; 95% CI, 13.1-25.4), consultation rate and overall score (β, 3.70; 95% CI, 1.4-5.9), symptom relief prescribing and management score (β, –7.30; 95% CI, –13.4 to –1.2), disease-specific minus symptom-relief prescribing rate and diagnosis score (β, 4.07; 95% CI, 1.2-7.0), and contraindicated prescribing and management score (relative risk, 0.89; 95% CI, 0.8-1.0).

†To facilitate interpretation of changes in the magnitude of the association between examination scores and practice outcomes over the first 4 to 7 years of practice, examination score outcome relationships were estimated for 3 time intervals of practice based on a categorization, for each physician, of the cumulative months in practice from the practice entry month. The interaction effects presented by intervals of years in practice were produced by a separate analysis to facilitate easier interpretation, in which the interactions between examination score and 2 dummy variables, representing practice years 3 to 4 and 5 to 7 relative to years 1 to 2 were estimated.

‡To test the hypothesis that the relationship between certification examination scores and practice outcomes would be attenuated with increasing time in practice, we tested the interaction between examination score and cumulative months in practice. Cumulative practice months, treated as a time-dependent covariate, were determined by counting each month that the physician billed the Québec health insurance agency (Régie de l'assurance maladie due Québec; RAMQ) for fee-for-service or salaried care for Québec medical care beneficiaries. *P* values are reported for each of the interaction terms (examination score multiplied by cumulative months in practice) that were estimated for each combination of outcome and examination score.

§Included phenylbutazone, dipyridamole, reserpine, disopyramide, clofibrate, methylphenidate, chlordiazepoxide, diazepam, clorazepate, flurazepam, clonazepam, clobazam, primidone, fluoxetine, phenelzine, tranylcypromine, moclobemide, amitriptyline, doxepin, imipramine, trimipramine, clomipramine, amoxapine, maprotiline, cyclobenzaprine, methocarbamol, pentazocine, meperidine, triazolam, and theophylline. Data expressed as relative risk of contraindicated prescribing per 1 SD increase in score.

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

have resulted in failing 16 additional physicians on the MCCQE licensing examination over 4 years. It would also have reduced the expected risk of contraindicated prescriptions for elderly patients seen by these physicians by approximately 42% (from 4.7% for these low-scoring physicians to 3.3% for an average physician). A reduction in risk of this magnitude is equivalent to or greater than that reported for the most effective form of continuing medical education for physicians with high rates of inappropriate prescribing.[43] To minimize the risk of adverse outcomes for the population, passing standards may need to be established for individual components of the examination, such as drug knowledge, diagnosis, and management.

Consultation rates increased linearly with examination score. Physicians tend to report higher referral rates in clinical areas in which they felt more competent.[33,34] More competent physicians may be more aware of their limitations. However, most examinations do

Table 4. Association Between Medical Council of Canada Licensing Examination Scores and Practice Performance in the First 4 to 7 Years of Practice

			Change in Outcome per SD Increase in Score					
			All Practice Years		Practice Year†			
Outcome	Mean (SD) Rate per 1000 Patients	Type of Certification Examination Score	β (95% CI)*	*P* Value	1-2	3-4	5-7	*P* Value for Interaction‡
Preventive care								
Mammography screening rate	122.6 (140.9)	Overall	16.81 (8.7 to 24.9)	<.001	15.2	21.5	15.3	.02
		Prevention	8.10 (–1.0 to 17.2)	.08	8.7	4.7	9.5	.17
		Clinical assessment	11.54 (5.5 to 17.6)	.002	10.5	14.5	11.1	.001
Coordination of care								
Continuity of care, % of visits	28.3 (11.0)	Overall	0.2 (–0.4 to 0.8)	.52	0.5	0.1	–0.1	.48
		Management	0.01 (–0.6 to 0.6)	.96	0.2	0.1	–0.3	.26
Resource use								
Consultation rate	77.76 (53.5)	Overall	4.93 (2.1 to 7.8)	<.001	4.9	5.4	5.0	.26
Acute and chronic disease management								
Symptom-relief prescribing rate among elderly patients	149.8 (118.7)	Overall	–7.15 (–15.0 to 1.0)	.08	–6.1	–9.3	–6.8	.36
		Diagnosis	–5.1 (–12.0 to 2.3)	.18	–3.2	–7.3	–5.7	.87
		Management	–7.0 (–14.0 to 2.2)	.15	–5.3	–6.1	–6.5	.51
		Drug knowledge	–6.99 (–14.0 to 0.4)	.06	–5.5	–8.3	–8.0	.61
Disease-specific prescription rate minus symptom-relief prescription rate	–2.11 (75.03)	Overall	4.83 (0.9 to 8.8)	.01	3.2	5.4	6.6	.30
		Diagnosis	3.80 (0.3 to 7.3)	.03	3.8	4.1	3.1	.69
		Management	3.0 (–1.0 to 6.9)	.15	1.2	3.8	4.6	.27
		Drug knowledge	3.5 (0 to 7.1)	.05	1.8	4.7	5.0	.50
Contraindicated prescribing among elderly patients§	32.58 (44.1)	Overall	0.93 (0.82 to 1.05)	.21	0.93	0.91	0.94	.44
		Diagnosis	0.97 (0.86 to 1.10)	.67	0.99	0.95	0.99	.75
		Management	0.91 (0.80 to 1.03)	.13	0.91	0.90	0.90	.54
		Drug knowledge	0.88 (0.77 to 1.0)	.05	0.88	0.87	0.88	.55

*The regression coefficient β represents the estimated change in the rate or value of the practice outcome per SD increase in score in the first 4 to 7 years in practice with a 95% confidence interval (CI). In these overall models, the interaction term to test the potential modification of the magnitude of the effect between practice outcome and examination scores in relationship with the number of months in practice is not included. In instances in which there was a significant interaction between examination score and months in practice (ie, mammography screening rate), the estimates for each category of years in practice provide a more appropriate estimate of the effects. Each β was estimated by a multivariate regression model within a generalized estimating equation framework, in which physician was the unit of analysis and annual assessment of outcome rates/values, were represented as repeated measurements for each physician. Observations were weighted by the logarithm of each physician's annual practice size. The estimate of the examination score, practice outcome relationship was adjusted for differences in annual practice case-mix including age and sex structure, socioeconomic status, geographic access to health care, comorbidity, and propensity to use health care services based on data for individual practice patients in the year prior to outcome assessment. The only exception was for contraindicated prescribing in which practice size was used to weight estimated regression coefficients and medical school was included, but practice case-mix covariates were not included as these attributes of the practice population would rarely justify contraindicated prescribing in the elderly. When medical school was excluded from the regression models, the magnitude of the association between examination scores and practice outcomes increased because some medical schools had systematically lower scores than others. If the analysis were based on usual practice, in which pass-fail decisions are made irrespective of medical school, the overall predictive relationship between examination scores and outcomes would have been higher: mammography screening and overall score (β, 18.3; 95% CI, 10.3-26.3), consultation rate and overall score (β, 5.13; 95% CI, 2.3-7.9), symptom relief prescribing and drug knowledge score (β, –8.36; 95% CI, –15.8 to –0.9), disease-specific minus symptom-relief prescribing rate and diagnosis score (β, 3.39; 95% CI, –0.4 to 6.8), and contraindicated prescribing and drug knowledge score (relative risk, 0.85; 95% CI, 0.7-0.9).

†To facilitate interpretation of changes in the magnitude of the association between examination scores and practice outcomes over the first 4 to 7 years of practice, examination score outcome relationships were estimated for 3 time intervals of practice based on a categorization, for each physician, of the cumulative months in practice from the practice entry month. The interaction effects presented by intervals of years in practice were produced by a separate analysis to facilitate easier interpretation, in which the interactions between examination score and 2 dummy variables, representing practice years 3 to 4 and 5 to 7 relative to years 1 to 2 were estimated.

‡To test the hypothesis that the relationship between certification examination scores and practice outcomes would be attenuated with increasing time in practice, we tested the interaction between examination score and cumulative months in practice. Cumulative practice months, treated as a time-dependent covariate, were determined by counting each month that the physician billed the Québec health insurance agency (Régie de l'assurance maladie due Québec; RAMQ) for fee-for-service or salaried care for Québec medical care beneficiaries. *P* values are reported for each of the interaction terms (examination score multiplied by cumulative months in practice) that were estimated for each combination of outcome and examination score.

§Included phenylbutazone, dipyridamole, reserpine, disopyramide, clofibrate, methylphenidate, chlordiazepoxide, diazepam, clorazepate, flurazepam, clonazepam, clobazam, primidone, fluoxetine, phenelzine, tranylcypromine, moclobemide, amitriptyline, doxepin, imipramine, trimipramine, clomipramine, amoxapine, maprotiline, cyclobenzaprine, methocarbamol, pentazocine, meperidine, triazolam, and theophylline. Data expressed as relative risk of contraindicated prescribing per 1 SD increase in score.

©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020

not test whether an individual knows when to refer. This may be an important area for test development, as the risk of outcomes related to medical errors has been shown to increase when physicians practice beyond their areas of competence.[44]

The finding that examinations taken at the end of medical school were also predictive of future practice may be useful for residency program directors, who could use this information to target learning opportunities to areas of deficiency. Medical school educators may also be able to identify persons more likely to experience difficulty in practice earlier in training, as scores on medical school examinations are strongly correlated with scores on licensing examinations.[45]

Future research should investigate methods of establishing outcomes-based passing scores, and develop more refined measures of quality of care in disease-specific populations. Longitudinal follow-up studies should be conducted in multijurisdictional cohorts of medical graduates to explore the aspects of admissions, undergraduate, and postgraduate training that may influence practice.

Author Contributions: *Study concept and design:* Tamblyn, Abrahamowicz, Dauphinee.
Acquisition of data: Tamblyn, Dauphinee, Grand'Maison, Brailovsky.
Analysis and interpretation of data: Tamblyn, Abrahamowicz, Hanley, Norcini, Girard.
Drafting of the manuscript: Tamblyn, Abrahamowicz, Dauphinee.
Critical revision of the manuscript for important intellectual content: Tamblyn, Abrahamowicz, Dauphinee, Hanley, Norcini, Girard, Grand'Maison, Brailovsky.
Statistical expertise: Tamblyn, Abrahamowicz, Hanley, Norcini, Girard.
Obtained funding: Tamblyn, Abrahamowicz, Dauphinee.
Administrative, technical, or material support: Tamblyn, Dauphinee, Girard.
Study supervision: Tamblyn.
Funding/Support: Research support was provided by the Canadian Institutes of Health Research, and the Fonds de Recherche en Santé du Québec. In addition, Dr Tamblyn and Dr Abrahamowicz are medical scientists at the Canadian Institutes of Health Research.
Acknowledgment: We thank Tim Wood, PhD, and André Phillipe Boulais, MSC, of the Medical Council of Canada for their expert assistance in data retrieval, linkage, and score reclassification; Joelle Lescop, MD, MPH, of the Québec College of Physicians who made this study possible; M. Jacques Barry, MBA, of the Régie de l'assurance maladie du Québec for his assistance in retrieving the health service data and responding to our many questions; and M. Jimmy Fragos, BSc, for his expert assistance in database management and variable creation.

REFERENCES

1. Irvine D. The performance of doctors, I: professionalism and self-regulation in a changing world. *BMJ.* 1997;314:1540-1542.
2. Tamblyn RM. Is the public being protected? prevention of suboptimal medical practice through training programs and credentialing examinations. *Eval Health Prof.* 1994;17:198-221.
3. Ayanian JZ, Hauptman PJ, Guadagnoli E, et al. Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction. *N Engl J Med.* 1994;331:1136-1142.
4. Edep ME, Shah NB, Tateo IM, Massie BM. Differences between primary care physicians and cardiologists in management of congestive heart failure. *J Am Coll Cardiol.* 1997;30:518-526.
5. Norcini J, Lipner RS, Benson JA Jr, Webster GD. An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. *Ann Intern Med.* 1985;102:385-389.
6. Jollis JG, DeLong ER, Peterson ED, et al. Outcome of acute myocardial infarction according to the specialty of the admitting physician. *N Engl J Med.* 1996; 335:1880-1887.
7. Weiner JP, Parente ST, Garnick DW, Fowles J, Lawthers AG, Palmer HR. Variation in office-based quality: a claims-based profile of care provided to Medicare patients with diabetes. *JAMA.* 1995;273:1503-1508.
8. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Wenrich MD. Predictive validity of certification by the American Board of Internal Medicine. *Ann Intern Med.* 1989;110:719-726.
9. McAuley RG, Paul WM, Morrison GH, Beckett RF, Goldsmith CH. Five-year results of the peer assessment program of the College of Physicians and Surgeons of Ontario. *CMAJ.* 1990;143:1193-1199.
10. Murata PJ, Li J. Relationship between Pap smear performance and physician ordering a mammogram. *J Fam Pract.* 1992;35:644-648.
11. Nicolucci A, Cavaliere D, Scorpiglione N, et al. A comprehensive assessment of the avoidability of long-term complications of diabetes. *Diabetes Care.* 1996; 19:927-932.
12. Legorreta AP, Christian-Herman J, O'Connor RD, et al. Compliance with national asthma management guidelines and specialty care. *Arch Intern Med.* 1998; 158:457-464.
13. Wilcox SM, Himmelstein DU, Woolhandler S. Inappropriate drug prescribing for the community-dwelling elderly. *JAMA.* 1994;272:292-296.
14. Tamblyn R, Abrahamowicz M, Brailovsky C, et al. Association between licensing examination scores and resource use and quality of care in primary care practice. *JAMA.* 1998;280:989-996.
15. *Statistiques Annuelles.* Québec, Québec: Regie de l'assurance-maladie du Québec; 1995.
16. Grand'Maison P, Lescop J, Rainsberry P, Brailovsky CA. Large-scale use of an objective, structured clinical examination for licensing family physicians. *CMAJ.* 1992;146:1735-1740.
17. *Information Pamphlet on the Qualifying Examination.* Ottawa, Ontario: Medical Council of Canada; 1991.
18. Livingston SA, Zieky M. *Passing Scores.* Princeton, NJ: Educational Testing Service; 1989.
19. Tamblyn RM, Lavoie G, Petrella L, Monette J. Use of prescription claims databases in pharmacoepidemiological research. *J Clin Epidemiol.* 1995;48:999-1009.
20. Levy AR, Tamblyn RFD, McLeod P, Hanley J. Coding accuracy of hospital discharge data for elderly survivors of myocardial infarction. *Can J Cardiol.* 1999; 15:1277-1282.
21. Wilchesky M, Tamblyn RM. Predictive values of diagnoses in medical services claims. *J Clin Pharmacol.* 2001;220(suppl).
22. Wilkins K. Use of postal codes and addresses in the analysis of health data. *Health Rep.* 1993;5:157-177.
23. Salive ME, Guralnik JM, Brock D. Preventive services for breast and cervical cancer in US office-based practices. *Prev Med.* 1996;25:561-568.
24. Bindman AB, Grumbach K, Osmond D, et al. Preventable hospitalizations and access to health care. *JAMA.* 1995;274:305-311.
25. Shea S, Misra D, Ehrlich MH, Field L, Francis CK. Predisposing factors for severe, uncontrolled hypertension in an inner-city minority population. *N Engl J Med.* 1992;327:776-781.
26. Wasson JH, Sauvigne AE, Mogielnicki RP, et al. Continuity of outpatient medical care in elderly men: a randomized trial. *JAMA.* 1984;252:2413-2417.
27. Lambrew JM, DeFriese GH, Carey TS, Ricketts TC, Biddle AK. The effects of having a regular doctor on access to primary care. *Med Care.* 1996;34:138-154.
28. Hulscher ME, van Drenth BB, Mokkink HG, et al. Barriers to preventive care in general practice. *Br J Gen Pract.* 1997;47:711-714.
29. Safran DG, Montgomery JE, Chang H, Murphy J, Rogers H. Switching doctors: predictors of voluntary disenrollment for primary care physician's practice. *J Fam Pract.* 2001;50:130-136.
30. McGavock H, Wilson-Davis K, Niblock RWF. Unsuspected patterns of drug utilization revealed by interrogation of a regional general practitioner prescribing database. *Pharmacoepidemiol Drug Saf.* 1992;1: 73-80.
31. Leape LL, Bates DW, Cullen DJ, et al. Systems analysis of adverse drug events. *JAMA.* 1995;274:35-43.
32. Franks P, Clancy CM. Gatekeeping revisited: protecting patients from overtreatment. *N Engl J Med.* 1992;327:424-429.
33. Reynolds GA, Chitnis JG, Roland MO. General practitioner outpatient referrals: do good doctors refer more patients to hospital? *BMJ.* 1991;302:1250-1252.
34. Calman NS, Hyman RB, Licht W. Variability in consultation rates and practitioner level of diagnostic certainty. *J Fam Pract.* 1992;35:31-38.
35. Beers MH, Ouslander JG, Rollingher I, Reuben DB, Brooks J, Beck JC. Explicit criteria for determining inappropriate medication use in nursing home residents. *Arch Intern Med.* 1991;151:1825-1832.
36. McLeod PJ, Huang AR, Tamblyn RM, Gayton DC. Defining inappropriate practices in prescribing for elderly people. *CMAJ.* 1997;156:385-391.
37. Starfield B, Weiner J, Mumford L, Steinwachs D. Ambulatory care groups: a categorization of diagnoses for research and management. *Health Serv Res.* 1991;26:53-74.
38. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with *ICD-9-CM* adminstrative databases. *J Clin Epidemiol.* 1992;45:613-619.
39. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continous outcomes. *Biometrics.* 1986; 42:121-130.
40. Public trust and accountability for clinical performance: lessons from the national press reportage of the Bristol hearing. *J Eval Clin Pract.* 1999;5:335-342.
41. Southgate L, Hays RB, Norcini J, et al. Setting performance standards for medical practice: a theoretical framework. *Med Educ.* 2001;35:474-481.
42. Norcini JJ Jr. Standards and reliability in evaluation. *Acad Med.* 1999;74:1088-1090.
43. Schaffner W, Ray WA, Federspiel CF, Miller WO. Improving antibiotic prescribing in office practice. *JAMA.* 1983;250:1728-1732.
44. Leape LL, Brennan TA, Laird N, et al. Nature of adverse events in hospitalized patients. *N Engl J Med.* 1991;324:377-384.
45. Veloski JJ, Hojat M, Gonnella JS. The validity of Part III of the National Board Examination. *Proc Annu Conf Res Med Educ.* 1987;26:54-59.

 ©2002 American Medical Association. All rights reserved.

Downloaded From: https://jamanetwork.com/ by a McGill University Libraries User on 12/20/2020