

# Performance assessment

## *Family physicians in Montreal meet the mark!*

François Goulet, MD, MA André Jacques, MD Robert Gagnon, MSC Denis Bourbeau, MD  
Denis Laberge, MD Jacques Melanson, MD Claude Ménard, MD Pierre Racette, MD Raymond Rivest, MD

### ABSTRACT

**OBJECTIVE** To assess the clinical performance of a representative non-volunteer sample of family physicians in metropolitan Montreal, Que.

**DESIGN** Assessment of clinical performance was based on inspection visits to offices, peer review of medical records, and chart-stimulated recall interviews. The procedure was the one usually followed by the Professional Inspection Committee of the Collège des médecins du Québec.

**SETTING** Family physicians' practices in metropolitan Montreal.

**PARTICIPANTS** One hundred randomly selected family physicians.

**INTERVENTIONS** For each physician, 30 randomly chosen patient charts with data on three to five previous visits were reviewed using explicit criteria and a standard scale using global scores from 1 to 5 (unacceptable to excellent).

**MAIN OUTCOME MEASURES** Scores were assigned for office practices; record keeping; number of continuing medical education (CME) activities; and quality of clinical performance assessed in terms of investigation plan, diagnostic accuracy, treatment plan, and relevance of care.

**RESULTS** Overall performance was judged to be good to excellent for 98% of physicians in their private practices; for 90% of physicians concerning CME activities; for 94% of physicians concerning their clinical performance in terms of quality of care; and for 75% of physicians as to record keeping. There was a link between record keeping and quality of care as well as between the number of CME activities and quality of care.

**CONCLUSION** The overall clinical performance of family physicians in the greater Montreal region is excellent.

### RÉSUMÉ

**OBJECTIF** Évaluer le rendement clinique d'un échantillon représentatif de médecins de famille sur une base non volontaire dans le Grand Montréal, au Québec.

**CONCEPTION** L'évaluation du rendement clinique se fondait sur des visites d'inspection dans les cabinets, la revue critique par des pairs des dossiers médicaux et des entrevues de rappel stimulées par les fiches médicales. Les modalités sont celles utilisées par le Comité d'inspection professionnelle du Collège des médecins du Québec.

**CONTEXTE** Des pratiques de médecins de famille dans le Grand Montréal.

**PARTICIPANTS** Cent médecins de famille choisis au hasard.

**INTERVENTIONS** Pour chaque médecin, 30 dossiers de patients choisis au hasard portant des données sur les trois à cinq dernières visites ont fait l'objet d'un examen en fonction de critères explicites et d'un barème standard utilisant des cotes globales de 1 à 5 (d'inacceptable à excellent).

**PRINCIPALES MESURES DES RÉSULTATS** Des cotes étaient accordées pour les pratiques de bureau; la tenue des dossiers; le nombre d'activités de formation médicale continue (FMC); et la qualité du rendement clinique évaluée en fonction du plan d'investigation, de l'exactitude du diagnostic, du plan thérapeutique et de la pertinence des soins.

**RÉSULTATS** Le rendement général était jugé de bon à excellent pour 98% des médecins dans leurs pratiques privées; pour 90% des médecins concernant leurs activités de FMC; pour 94% des médecins concernant leur rendement clinique en terme de qualité des soins; et pour 75% des médecins quant à leur tenue de dossiers. Il existait un rapport entre la tenue de dossiers et la qualité des soins ainsi qu'entre le nombre d'activités de FMC et la qualité des soins.

**CONCLUSION** Le rendement clinique global des médecins de famille dans la région du Grand Montréal est excellent.

*This article has been peer reviewed.*

*Cet article a fait l'objet d'une évaluation externe.*

*Can Fam Physician 2002;48:1337-1344.*

## RESEARCH

### ..... Performance assessment

**S**ince the early 1980s, medical licensing authorities have become increasingly interested in monitoring and maintaining physicians' performance. Peer assessment programs based mainly on studying patients' medical records are practical because clinical performance can be evaluated in terms of clinical management, diagnosis, use of resources, treatment approaches, and follow-up care.<sup>1-5</sup> In addition, these programs allow identified shortcomings to be corrected and, when combined with refresher training programs, contribute toward maintaining performance.

Some licensing authorities have developed a multifaceted approach that integrates formal continuing medical education (CME) activities, selected readings, quality assessment programs, and, in one case, surveys of patients and colleagues; others have developed peer assessment programs.<sup>6-15</sup>

Performance assessment programs through chart audit have been judged valid, within certain limits.<sup>3,16-18</sup> Interrater reliability is still under study and has shown only mixed results in past reviews.<sup>19,20</sup> To increase the validity of evaluations, chart-stimulated recall of physicians visited often allows researchers to gather data not included in clinical notes.<sup>21</sup>

Since 1973, the Collège des médecins du Québec (CMQ), the medical licensing authority, has been required by provincial law to assess the practice of Quebec physicians on a non-voluntary basis.<sup>22</sup> The Professional Inspection Committee (PIC) analyzes clinical performance during individual office visits to physicians. Their reports are based on peer assessment using chart audits and chart-stimulated recall interviews with the physicians assessed. Individual office visits are planned in relation to clinical or administrative practice profiles (using indicators) identifying physicians who could have problems.

In many cases where physicians were identified as having potentially unsatisfactory profiles, they were ultimately judged to be having problems. In each case, the PIC judges shortcomings in relation to the expected quality of professional practice among all Quebec physicians.

To better estimate what this quality is, the PIC set out to assess the professional practices of a randomly

.....  
**Dr Goulet** is Assistant Director; **Dr Jacques** is Director; **Dr Bourbeau** is a former Director; **Dr Laberge**, **Dr Melanson**, **Dr Ménard**, **Dr Racette**, and **Dr Rivest** are Inspector-Investigators; and **Mr Gagnon** is a Research Associate in the Practice Enhancement Division of the Collège des médecins du Québec in Montreal, Que.

selected group of 100 non-volunteer family physicians in metropolitan Montreal, Que. This article presents the results of that benchmark assessment.

## METHODS

### Sample selection

All 3767 family physicians from the CMQ membership list who practise either entirely or partly in private offices in greater Montreal (Montreal, Laval, and the South Shore) were targeted as the study population. This geographical group was chosen to limit the cost of professional inspection visits. In this paper, all general practitioners with or without certification by the College of Family Physicians of Canada are referred to as family physicians. A group of 100 physicians was randomly selected by a computer from the list of all eligible physicians in greater Montreal. A stratification criterion was used to ensure that the final sample contained similar numbers of physicians with fewer than 15 years in practice and physicians with more than 15 years in practice.

### Assessment process

Professional inspection visits were conducted over 1 year (1996-1997) by five investigators (one internist, one general surgeon, one urologist, and two family physicians) who worked exclusively for the CMQ in the Professional Inspection Division. They followed a 6-month training program in peer review and had an average of 10 years' (1 to 17 years) experience at full-time performance assessment.

Each visit allowed investigators to study at least 30 patient charts randomly chosen from a list of patients who reflected the physician's practice population and had consulted the physician investigated during 1 or 2 typical days of practice in the past 3 months. Each chart was reviewed to acquire data on three to five previous office visits. This number of visits is practical and offers good face validity. Because no specific health conditions were selected, the review was done for all conditions actually mentioned in the charts. This evaluation allowed for both cross-sectional (different patients) and longitudinal (same patient over time) study.

After reviewing the records, a 1-hour chart-stimulated recall interview was conducted where charts reviewed by the inspector were discussed with the physician investigated. They discussed how the physician went about gathering data, diagnosing, solving problems, managing patients' problems, using resources, and keeping records. This discussion

enabled investigators to validate the information gathered from records and to complete the evaluation.

Private office practices were evaluated using explicit criteria according to practice guidelines and regulations published by the CMQ: physical area (clean, in order, well lit, rest rooms available, appropriate furniture, and respectful of privacy), medical apparatus (stethoscope, sphygmomanometer, oto-ophthalmoscope, scale, glucometer), prevention of infection (cleaning of instruments, presence or absence of a sterilizer, management of biochemical garbage), and office management (management of laboratory tests' follow up, registration of patients).

The score for record keeping was based on legibility, organization, and documentation of several items: name, sex, address, and date of birth of the patient; important medical information (history, physical examination, diagnoses, laboratory and radiologic tests, medications, counseling, follow up, family and social history, allergies).

To assess the number of CME activities in which physicians were involved, the inspector asked physicians for a list of educational activities they participated in over the past year. Inspectors checked the number of hours spent on CME, the relevance of the programs, and physicians' subscriptions to medical journals.

Quality of care was evaluated using four parameters: investigation plan, accuracy of diagnosis, treatment plan, and relevance of care. The parameters of care were evaluated in relation to health problems that are frequently encountered in family medicine (high blood pressure, diabetes, low back pain, chronic pulmonary obstructive disease, asthma, the periodic health examination, etc). A panel of family physicians and researchers modified the explicit criteria for those problems based on evidence from the scientific literature (when possible).<sup>23,24</sup> Examples of two health problems frequently encountered in family medicine are presented in **Figures 1** and **2**.

All assessments were conducted using a standard grid listing the criteria; it included a scale with scores from 1 to 5 (1—unacceptable, 2—inadequate, 3—good, 4—very good, 5—excellent). For each physician, a global score was assigned to each variable on the basis of all information obtained from chart reviews and chart-stimulated recall interviews.

The Quebec Professional Code allows the PIC to evaluate any physician practising medicine in Quebec, and this applies to all places of practice. Thus, this study was done inside the regular activities of the PIC, and all physicians selected and visited were subjected to the same procedures usually

followed under other PIC programs. Thus no physician could refuse to participate.

### Data analysis

As proposed by Norton and colleagues,<sup>14</sup> we used a binary scoring system in the analyses by grouping scores of 1 and 2 under the term unsatisfactory and scores 3 to 5 under the term satisfactory.<sup>14</sup> Analyses were conducted using the SPSS statistics package. Differences between subgroups were studied using the  $\chi^2$  test and *t* test for means. The  $\phi$  coefficient derived from calculation of the  $\chi^2$  was used for dichotomous data, and the Spearman rank correlation was used for ordinal scales and Pearson correlation for continuous scales.

A sample size of 100 physicians was chosen in order to obtain a 15% confidence interval (CI) for expected proportion of satisfactory scores of 80%. All CIs are at the 95% level, and all statistical analyses are two-tailed and significant at the level of  $P \leq 5\%$ .

## RESULTS

Comparisons between family physicians selected (n=100) and all physicians in the target population (n=3767) show no statistically significant differences on any sociodemographic variables (**Table 1**).

**Table 2** and **Figure 3** show results of assessments for each variable measured. Thus, high proportions of physicians received scores for quality of professional practice in the range good to excellent: investigation plan (85%), diagnostic accuracy (97%), relevance of care (98%), and treatment plan (96%).

**Table 1. Demographic data:** Average age of physicians in the study was 47.6 years (standard deviation 8.7) and of Collège des médecins du Québec members in Montreal was 47.6 years ( $P = 1.0$ ).

PHYSICIANS' CHARACTERISTICS	STUDY SAMPLE (N = 100) (%)	COLLÈGE DES MÉDECINS DU QUÉBEC (METROPOLITAN AREA) AVERAGE (N= 3767) (%)	P VALUE
<b>SEX</b>			.31
Male	57	62	
Female	43	38	
<b>LICENCE TO PRACTISE</b>			.43
1980	48	44	
Before 1980	52	56	
<b>MEDICAL DEGREE</b>			.24
Quebec universities	73	78	
Other universities	27	22	

**RESEARCH**  
 .....  
 Performance assessment

**Table 2. Ratings of practice performance among Montreal family physicians**

PRACTICE CHARACTERISTIC	1 UNACCEPTABLE N (%)	2 INSUFFICIENT N (%)	3 GOOD N (%)	4 VERY GOOD N (%)	5 EXCELLENT N (%)	NOT EVALUATED	SATISFACTORY PERFORMANCE* (%)	95% CONFIDENCE INTERVAL
Office practices	0	2 (2.1)	6 (6.3)	34 (35.4)	54 (56.3)	4	97.9	95-100
Record keeping	4 (4)	21 (21)	32 (32)	27 (27)	16 (16)	0	75.0	66.5-83.5
Continuing medical education	5 (5.1)	5 (5.1)	27 (27.3)	31 (31.3)	31 (31.3)	1	89.9	84-95.8
Investigation plan	0	15 (15.3)	28 (28.6)	36 (36.2)	19 (19.4)	2	84.7	77.6- 91.8
Diagnostic accuracy	0	3 (3.1)	27 (27.8)	39 (40.2)	28 (28.9)	3	96.9	93.5-100
Treatment plan	0	4 (4.1)	25 (25.8)	42 (43.3)	26 (26.8)	3	95.9	91.9-99.9
Relevance of care	0	2 (2)	27 (27.6)	36 (36.7)	33 (33.7)	2	98.0	95.2-100

\*Grouped scores for good, very good, and excellent.

**Figure 1. Health chart for periodic health examination of adults aged 45 to 64**

**Canadian Task Force on the Periodic Health Examination 1994**

FILE NUMBER \_\_\_\_\_ PATIENT'S NAME \_\_\_\_\_ SEX \_\_\_\_\_

YEAR/MONTH/DAY

AGE GROUP	45-49	50-54	55-59	60-64
-----------	-------	-------	-------	-------

**MEDICAL HISTORY AND PHYSICAL EXAMINATION**

Blood pressure  
 Skin examination  
 Breast examination

**IMMUNIZATION AND LABORATORY TESTS**

Diphtheria and tetanus every 10 years  
 Influenza (annually)  
 Cholesterol (men)  
 Mammography  
 Cervical cytology\*

**COUNSELING**

Smoking  
 Nutrition  
 Physical activity  
 Lifestyle (sun and alcohol)  
 Accidents  
 Oral hygiene  
 Sexuality  
 Hormone therapy

**SPECIFIC PROBLEMS**

\*Annual testing recommended following initiation of sexual activity or after 18. If the first two tests are normal, may be repeated every 3 years. May be performed more frequently for women at risk.

■ Recommended if at risk

ALL RIGHTS RESERVED

Collège des médecins du Québec  
 COLLEGE OF FAMILY PHYSICIANS OF CANADA, QUEBEC SECTION  
 Health Canada

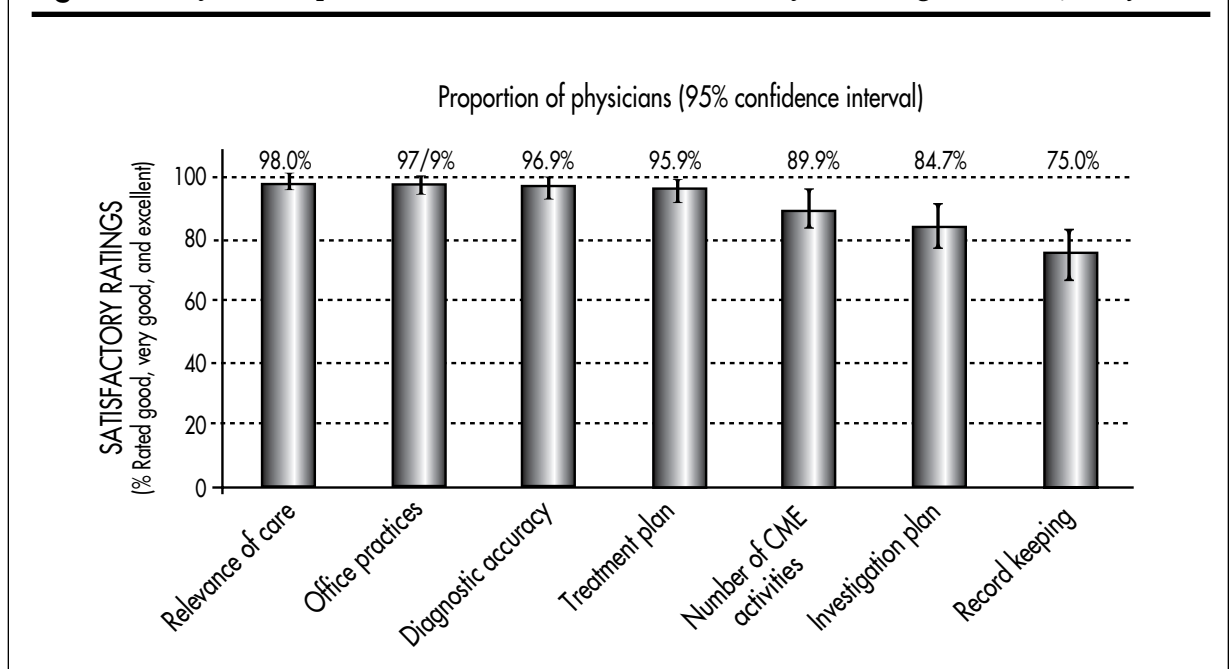
**Figure 2. Diagnosis and treatment of low back pain**

FILE NUMBER \_\_\_\_\_ PATIENT'S NAME \_\_\_\_\_ SEX \_\_\_\_\_

**QUESTIONNAIRE**

	PRESENCE	ABSENCE
Symptoms (length of time present)		
Factors aggravating pain		
Reference of pain		
<b>PHYSICAL EXAMINATION</b>		
Anterior flexion of spine (spinal motion)		
Neurologic examination		
Knee reflexes		
Ankle reflexes		
Walking on toes		
Walking on heels		
Great toe dorsiflexion strength		
Straight leg-raising test		
<b>TREATMENT (IF NO SERIOUS CAUSE SUSPECTED)</b>		
Analgesia (acetaminophen, nonsteroidal anti-inflammatory drugs, etc)		
Counseling (rest)		
Ice or heat		

**Figure 3. Physicians' performance on variables under study: Percentage rated satisfactory**



Overall, 94% of physicians received good-to-excellent scores for the quality of their practice. Amount of CME was judged to be good to excellent in 92% of cases. Office practices were judged to be good to excellent for 98% of physicians, but only 75% of

physicians were judged to have good record-keeping practices.

Associations between quality of care variables and characteristics of physicians were analyzed (Table 3). Demographic variables and variables related to

## RESEARCH

### Performance assessment

quality of care were significantly but weakly correlated. Positive associations between investigation plan and age ( $r = 0.20$ ;  $P < .05$ ) and being a female physician ( $r = 0.26$ ;  $P < .05$ ) were observed. There was a relation between record keeping and quality of care ( $r = 0.33$  for diagnostic accuracy and  $r = 0.57$  for investigation plan;  $P \leq .001$ ) and between amount of CME activity and quality of care ( $r = 0.33$ ,  $P \leq .001$  for diagnostic accuracy and  $r = 0.23$ ,  $P \leq .01$  for investigation plan).

**Table 3. Correlations between physicians' characteristics and three variables measuring quality of care**

PHYSICIANS' CHARACTERISTICS	DIAGNOSTIC ACCURACY	INVESTIGATION PLAN	TREATMENT PLAN
Age	0.06	0.20*	0.15
Sex	0.16	0.26*	0.19
University where physician received MD degree	0.20	0.10	0.23
Number of office visits weekly	0.22*	0.01	0.15
Amount of continuing medical education	0.33 <sup>†</sup>	0.23 <sup>‡</sup>	0.10
Record keeping	0.33 <sup>†</sup>	0.57 <sup>†</sup>	0.14

\* $P < .05$ .

<sup>†</sup> $P < .001$ .

<sup>‡</sup> $P < .01$ .

## DISCUSSION

This study of practice performance is innovative because it used non-volunteer recruitment of subjects and added stimulated chart recall to chart review. Researchers assessed quality of care, record keeping, and amount of CME activity in a random sample of family physicians practising in the Montreal region. The group is likely to be representative of family physicians in greater Montreal because its demographics are comparable to those of physicians on the general CMQ membership list who practise in metropolitan Montreal.

Diagnostic accuracy and treatment plans are key elements in physicians' clinical procedures. Because almost all physicians in this study scored in the good-to-excellent range for diagnostic accuracy and their treatment plans, these results reassure the licensing authority and the general public about the performance of family physicians in the Montreal region. The results obtained in this region of Quebec are more or less similar to those reported by McAuley and Henderson<sup>10</sup> or by

Norman and associates<sup>11</sup> in Ontario. The satisfactory score earned by 90% of physicians for amount of CME activity corresponds to figures presented previously.<sup>25</sup>

Results of this study differ on some points from those of Norton and colleagues<sup>14</sup> because no significant differences were found between physicians' performance and the university where their medical degree was earned or the number of patients seen weekly. Age and sex were only weakly correlated.<sup>14</sup> The study did not assess certification by the College of Family Physicians of Canada because it was not mandatory in Quebec before 1988, and very few older physicians had this certification.

It is unsurprising that we found some significant associations between record keeping and three of the four aspects of quality of professional practice. Obviously, physicians' professional competence influences the quality of the elements found in a patient's records, yet quality of record keeping does not necessarily reflect a physician's performance. This phenomenon highlights the limitations of using patients' medical records as an evaluation tool. Nevertheless, the chart-stimulated recall interview enabled investigators to evaluate physicians' clinical management. To a certain extent, such an interview allows a line to be drawn between quality of practice and quality of record keeping. In addition, the link between amount of CME and the same three aspects of quality of practice should be explored further by licensing authorities before imposing a mandatory number of CME hours physicians must put in to maintain a valid licence to practise.

Interrater reliability could not be evaluated without compromising professional ethics. It would have been contrary to the usual practice of the PIC to conduct professional inspection visits with many inspectors or even to have them make these visits consecutively. Contrary to other peer assessment programs that use a large number of inspectors and reviewers, only five experienced inspectors, each making more than 40 visits yearly, participated in the study.<sup>16</sup> Their vast experience likely ensures greater reliability than would little experience in chart audits by a lot of other clinicians. Inspectors also used explicit criteria to facilitate their task. Each of the evaluators appeared to give the same average scores for each criterion. This leads us to believe that there is not much interrater variability in judgment. A College study is currently assessing the interrater reliability of PIC evaluations.

This study, like others assessing the quality of practice through peer review of patient records,

has two main limitations.<sup>10,14</sup> The first deals with the quality of records kept for each patient, which largely limits the elements that allow quality of practice to be assessed.<sup>26</sup> Chart-stimulated recall interviews help to minimize the importance of this limitation. So, for pragmatic reasons and to respect the usual procedures of the PIC, audiotaped or videotaped office consultations were not considered even if some studies showed good validity and feasibility.<sup>27,28</sup> The second limitation is that this type of review does not allow assessment of physicians' technical skills during physical examinations, communication skills, skills in minor surgical interventions, and patient satisfaction. Further, the small sample size used in this study, while providing a satisfactory overall assessment of physicians' performance, did not yield sufficient power for adequate subgroup analyses.

## CONCLUSION

The CMQ recently chose to make professional inspection visits and to conduct peer reviews of patient records as a tool for assessing physicians' professional practices. This study of family physicians enables the authors to reassure the general public and to congratulate family physicians in greater Montreal on the overall high quality of their practices. \*

### Acknowledgment

We thank Mrs Nicole Hanneman for her assistance in transcribing the document.

### Contributors

All authors participated in writing and revising this article and approved the final submission; only the last five made inspection visits to offices. **Dr Goulet** drafted the article and analyzed and interpreted data. **Mr Gagnon** analyzed and interpreted data and performed statistical analyses. **Dr Jacques** analyzed and interpreted data. **Dr Bourbeau** conceived of and designed the project. **Dr Laberge** conceived of and designed the project and acquired and interpreted data. **Dr Melanson, Dr Ménard, Dr Racette, and Dr Rivest** acquired data.

### Competing interests

None declared

**Correspondence to:** Dr François Goulet, Practice Enhancement Division, Collège des médecins du Québec, 2170 René Lévesque Blvd W, Montréal, QC H3H 2T8; e-mail [goulet.cmq@sympatico.ca](mailto:goulet.cmq@sympatico.ca)

### Editor's key points

- This study examined the quality of care delivered by Montreal family physicians as determined by the Professional Inspection Committee of the Collège des médecins du Québec.
- The study was innovative in that assessment was not voluntary and in that chart-stimulated recall was used as well as chart review.
- A high proportion of these physicians (85% to 98%) had good to excellent ratings on clinical performance (investigation and treatment), office practices, and continuing medical education. A lower rating (75% satisfactory) was seen for record keeping.
- There was a moderate association between good record keeping and amount of continuing medical education and quality of care.

### Points de repère du rédacteur

- Cette étude examinait la qualité des soins dispensés à Montréal par les médecins de famille conformément à la procédure du Comité d'inspection professionnelle du Collège des médecins du Québec.
- L'étude était novatrice en ce sens que l'évaluation n'était pas sur une base volontaire et qu'on s'est servi d'une entrevue de rappel stimulée par les fiches médicales en plus de l'examen des dossiers.
- Une grande proportion de ces médecins (de 85% à 98%) ont obtenu d'excellentes cotes pour le rendement clinique (investigation et thérapie), les pratiques de bureau et la formation médicale continue. Une cote plus faible (75%) a été observée pour la tenue des dossiers.
- Une association modérée se dégageait entre la bonne tenue de dossiers, la quantité de formation médicale continue et la qualité des soins.

### References

1. Langsley DG. Medical competence and performance assessment; a new era. *JAMA* 1991;266:977-80.
2. Dunn EV, Bass MJ, Williams JI, Borgiel AE, MacDonald P, Spasoff RA. Study of relation of continuing medical education to quality of family physicians' care. *J Med Educ* 1988;63(10):775-84.
3. Tugwell P, Dok C. Medical record review. In: Neufeld VR, Norman GR, editors. *Assessing clinical competence*. New York, NY: Springer Publishing Co; 1985. p. 142-82.
4. Wakefield DS, Helms CM, Helms L. The peer review process: the art of judgment. *J Healthcare Qual* 1995;17(3):11-6,51.
5. Borgiel AEM, Williams JI, Anderson GM, Bass MJ, Dunn EV, Lamont CT, et al. Assessing the quality of care in family physicians' practice. *Can Fam Physician* 1985;31:853-62.
6. Janes RD, Turner N. Reaccreditation in general practice; how New Zealand approaches the solution. *Can Fam Physician* 1995;41:1733-8.
7. Hall W, Violato C, Lewkonja R, Lockyer J, Fidler H, Toews J, et al. Assessment of physician performance in Alberta: the physician achievement review. *Can Med Assoc J* 1999;161:52-7.
8. Royal College of General Practitioners. *What sort of doctor? Assessing quality of care in general practice*. London, Engl: Royal College of General Practitioners; 1985.
9. Eliasson G, Berg L, Carlsson P, Lindstrom K, Bengtsson C. Facilitating quality improvement in primary health care by practice visiting. *Qual Health Care* 1998;7(1):48-54.

## RESEARCH

.....

### Performance assessment

10. McAuley RG, Henderson HW. Results of the peer assessment program of the College of Physicians and Surgeons of Ontario. *Can Med Assoc J* 1984;131(6):557-61.
11. Norman GR, Davis DA, Lamb S, Hanna E, Caulford P, Kaigas T. Competency assessment of primary care physicians as part of a peer review program. *JAMA* 1993;270(9):1046-51.
12. Page GG, Bates J, Dyer SM, Vincent DR, Bordage G, Jacques A, et al. Physician-assessment and physician-enhancement programs in Canada. *Can Med Assoc J* 1995;153(12):1723-8.
13. Norton PG, Dunn EV, Soberman L. Family practice in Ontario; how physician demographics affect practice patterns. *Can Fam Physician* 1994;40:249-56.
14. Norton PG, Dunn EV, Soberman L. What factors affect quality of care? Using the Peer Assessment Program in Ontario family practices. *Can Fam Physician* 1997;43:1739-44.
15. 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. *Can Med Assoc J* 1990;143(11):1193-9.
16. Rethans JJ, Martin E, Metsemakers J. To what extent do clinical notes by general practitioners reflect actual medical performance? A study using simulated patients. *Br J Gen Pract* 1994;44(381):153-6.
17. Wilson A, McDonald P. Comparison of patient questionnaire, medical record, and audio tape in assessment of health promotion in general practice consultations. *BMJ* 1994;309(6967):1483-5.
18. Norman GR, Painvin DA, Lindsay E, Rath D, Ragbeer M. Comprehensive assessment of clinical competence of family/general physicians using multiple measures. In: *Research in medical education: 1989 Proceedings of the twenty-eighth annual conference in conjunction with the 100th annual meeting of the Association of American Medical Colleges*; 1989 Oct 27-Nov 2; Washington, DC: Association of American Medical Colleges; 1989. p. 75-9.
19. Goldman RL. The reliability of peer assessments. A meta-analysis. *Eval Health Professions* 1994;17(1):3-21.
20. Smith MA, Atherly AJ, Kane RL, Pacala JT. Peer review of the quality of care: reliability and sources of variability for outcome and process assessments. *JAMA* 1997;278(19):1573-8.
21. Munger BS, Reinhart MA. Trial of multiple recertification methods. In: *Recertification for medical specialists*. Evanston, Ill: American Board of Medical Specialties; 1987. p. 71-88.
22. Québec (Province). *Professional code*. G.O. II 1973;105(26): 4949, art.107.
23. The Cochrane Collaboration. The Cochrane Library [database of systematic reviews on Internet, disk, and CD-ROM]. Oxford, Engl: Update Software; 1996. Updated quarterly. Available at: [www.update-software.com/cochrane.htm](http://www.update-software.com/cochrane.htm). Accessed 2002 April 12.
24. Canadian Task Force on the Periodic Health Examination. *The Canadian guide to clinical preventive health care*. Ottawa, Ont: Health Canada; 1994.
25. Goulet F, Gagnon RJ, Desrosiers G, Jacques A, Sindon A. Participation in CME activities; family physicians in practice for more than 25 years versus those in practice for less than 25 years. *Can Fam Physician* 1998;44:541-8.
26. Payne BC. The medical record as a basis for assessing physician competence. *Ann Intern Med* 1979;91(4):623-9.
27. Ram P, Van der Vleuten C, Rethans J-J, Grol R, Aretz K. Assessment of practicing family physicians: comparison of observation in multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med* 1999;74(1):62-9.
28. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-54.