What factors influence primary care physicians' charges for their services?

An exploratory study using standardized patients

Christel A. Woodward,*‡ PhD; Brian Hutchison,*†‡ MD, MSc;
Geoffrey R. Norman,* PhD; Judy A. Brown,* MA; Julia Abelson,* MSc

Abstract

Objective: To determine the extent of variation in physicians' charges for health care encounters with unannounced standardized patients and factors associated with the variation.

Design: Cross-sectional study.

Setting: Family practices open to new patients within 1 hour's drive of Hamilton, Ont.

Participants: A stratified random sample of 125 physicians who had responded to an earlier survey regarding preventive care were invited to participate. Of the 125, 44 (35.2%) declined to participate, and an additional 19 (15.2%) initially consented but later withdrew because they closed their practices to new patients. Sixty-two physicians thus participated in the study.

Intervention: Unannounced standardized patients posing as new patients to the practice visited study physicians' practices between September 1994 and August 1995, portraying 4 scenarios: 28-year-old woman, 52-year-old woman, 48-year-old man and 70-year-old man.

Outcome measures: Physician characteristics, encounter characteristics and charges made for services.

Results: The 62 physicians had 246 encounters with the standardized patients. Charges were made to the health insurance plan for services by 59 physicians for up to 4 encounters (215 encounters in all). Charges varied considerably both within and across patient scenarios. Time spent with the patient was an important factor predicting charges made ($p < 0.01$), although the effect of time spent on charges varied across scenarios ($p < 0.01$). Fee-for-service physicians charged more for their services than physicians who usually had alternative billing arrangements ($p < 0.01$). Female physicians charged more for their services than their male colleagues ($p = 0.03$). No relation was found between quality of preventive care and charges made ($p = 0.15$).

Conclusions: Physician-related factors are better able to account for the variability in charges for their services than patient-related factors. Physicians seeing comparable patients may earn much more or less than their colleagues because of differences in the services they provide and the way they apply the fee schedule. Quality-assurance techniques are likely needed to reduce the variability in charges seen and increase value for money spent in health care.

Résumé

Objectif : Déterminer l'importance de la variation des frais facturés par les médecins pour des consultations données à des patients normalisés non annoncés, ainsi que des facteurs associés à la variation.

Concept : Étude transversale.

Contexte : Cabinets de médecine familiale accueillant de nouveaux patients situés à moins d'une heure de route de Hamilton (Ont.).

Participants : Un échantillon aléatoire stratifié de 125 médecins qui avaient répondu à un sondage antérieur portant sur les soins préventifs ont été invités à

participer. Sur les 125, 44 (35,2 %) ont refusé de participer et 19 autres (15,2 %) y avaient consenti au début mais se sont retirés par la suite parce qu'ils ont cessé d'accueillir de nouveaux patients. Soixante-deux médecins ont donc participé à l'étude.

Intervention : Des patients normalisés non annoncés se faisant passer pour de nouveaux patients se sont rendus au cabinet des médecins participants entre septembre 1994 et août 1995. Ils représentaient 4 scénarios : femme de 28 ans, femme de 52 ans, homme de 48 ans et homme de 70 ans.

Mesures de résultats : Caractéristiques des médecins, caractéristiques de la consultation et frais facturés pour les services.

Résultats : Les 62 médecins ont donné 246 consultations aux patients normalisés. Cinquante-neuf médecins ont facturé au régime d'assurance-maladie des services pour jusqu'à 4 consultations (215 consultations au total). Les frais ont varié considérablement dans le même scénario et entre les scénarios. Le temps consacré aux patients a joué un rôle important dans la prévision des frais facturés ($p < 0,01$), même si l'effet du temps consacré sur les frais a varié entre les scénarios ($p < 0,01$). Les médecins rémunérés à l'acte ont facturé davantage pour leurs services que ceux qui utilisaient habituellement d'autres modes de facturation ($p < 0,01$). Les femmes médecins ont facturé davantage pour leurs services que leurs collègues masculins ($p = 0,03$). On n'a trouvé aucun lien entre la qualité des soins préventifs et les honoraires facturés ($p = 0,15$).

Conclusions : Les facteurs liés aux médecins expliquent davantage la variation des honoraires facturés pour leurs services que les facteurs liés aux patients. Les médecins qui accueillent des patients comparables peuvent gagner beaucoup plus ou moins que leurs collègues à cause de différences au niveau des services qu'ils fournissent et de leur façon d'appliquer la grille des honoraires. On a probablement besoin de techniques d'assurance de la qualité pour réduire la variabilité constatée dans les honoraires facturés et optimiser davantage les ressources consacrées aux soins de santé.

I t is generally assumed that since physicians in each Canadian province have a common fee schedule, their charges for seeing comparable patients would be highly similar. However, information about the extent to which variation occurs in physician charges or, indeed, whether such patients are offered similar services is lacking. Studies in which standardized patients[1,2] have visited community-based physicians and dentists have shown considerable variability in practice.[3–10] Given identical patient requests, health care providers vary widely in the time spent with the patient, the services provided or suggested, and the extent to which the services offered match appropriateness criteria developed by panels of expert clinicians or widely accepted guidelines for care provision.[3–10] These studies suggest that variation in the behaviour of individual health care providers may parallel the variability in service delivery seen in small-area variation studies.[11,12]

To our knowledge, only one study has examined the cost of office visits by standardized patients.[10] The study was conducted in northern New England, where no uniform fee schedule for physicians' services exists and where both general internists and family physicians deliver primary care. The authors found a 4-fold variation in charges for an initial visit for a "checkup" to establish ongoing care (mean US$53.54 [standard deviation (SD) US$20.18]). Although charges for services correlated with time spent, considerable unexplained variance remained.

In this article we examine the types of service offered and the extent of variation in costs of services provided during primary care visits by standardized patients to Ontario general practitioners and family physicians who used a uniform schedule of benefits paid by the provincial health insurance plan. We examine the extent to which factors related to the patient, the encounter and the physician account for differences in cost.

Methods

The methods for this study are described in detail in an accompanying paper (page 185). Briefly, we invited a stratified random sample of family physicians who practised in southern Ontario (within 1 hour's drive of Hamilton) to participate in a study in which 4 unannounced standardized patients visited their practices between September 1994 and August 1995. The physicians had responded to an earlier survey regarding preventive care practices and were accepting new patients.

The 4 standardized patient scenarios were a 48-year-old man, a 70-year-old man, a 28-year-old woman and a 52-year-old woman. Two people were trained to portray each scenario. Each physician was visited by one member of the patient pair. The standardized patients scheduled their own appointments with physicians.

The "patients" used bogus health insurance cards to allow unannounced entry into physicians' practices and payment of physicians for visits. Payment records were retrieved from the provincial health insurance plan, which was then reimbursed.

Physician and practice information was obtained in the initial preventive care survey. Data relating to the standardized patient–physician encounters were obtained from recall forms (e.g., amount of time spent) or derived from them (preventive care performance scores). Billing codes used, dollars expended and diagnostic information used for billing purposes were obtained from the Ontario Ministry of Health.

Descriptive analyses (e.g., frequency distributions and means) were initially used to examine the data and detect data entry errors. We examined bivariate associations using the statistics appropriate to the underlying data. To examine whether variability in billing could be explained by physician characteristics (sex, decade of graduation, usual remuneration type [fee-for-service or not], group practice [Yes/No], certification in family medicine [Yes/No] and problem-based undergraduate medical program [McMaster University, Hamilton, Ont.] [Yes/No]), we created dummy variables for each of these independent variables. They were entered into a linear regression equation along with characteristics of the care encounter (amount of time spent, preventive care performance score, standardized patient suspected or detected [Yes/No] and diagnostic tests ordered [Yes/No]), with cost of office-based services as the dependent variable. We used repeated-measure analysis of covariance (ANCOVA) to examine the effects of these variables across patients in predicting costs, while also examining the contribution of the patient scenarios and the physicians themselves.

The study was approved by the Ethics Review Committee of McMaster University's Faculty of Health Sciences.

Results

The characteristics of the 480 physicians who completed the initial preventive care survey and the subgroup of 251 physicians who had open practices and were eligible to participate in the standardized patient study differed significantly in only one variable: certified members of the College of Family Physicians of Canada were more likely than noncertificants to have closed practices and thus were less likely to be approached.

Of the 125 physicians (75 men and 50 women) approached about seeing standardized patients, 44 (35.2%) were not interested in participating, and 19 (15.2%) initially consented but later withdrew because they closed their practices to new patients. Sixty-two physicians thus participated in the study. No difference was noted in background characteristics between refusers and consenters except that refusers were significantly more likely to have graduated from medical school in the 1980s.

Most of the 62 physicians were in fee-for-service practice (82.2%), worked in groups (72.6%), were certified in family medicine (66.1%) and were male (64.5%). They reported seeing 4.8 patients per hour on average. Female physicians reported having a higher proportion of women patients than did male physicians (74% v. 52%) ($p < 0.001$).

Of the 62 physicians, 60 saw all 4 standardized patients. The billing data showed that 3 physicians with alternative funding arrangements had not billed for any of the patients seen. Of the 59 physicians who submitted charges, 43 had billed for all 4 standardized patients, 11 had billed for 3 patients, and 5 had billed for 2 patients. Thus, information on charges was available for 52 to 55 physicians for each scenario and for a total of 215 encounters.

Table 1 presents the characteristics of the encounters for each scenario. Physicians spent twice as much time with the 70-year-old man as with the 28-year-old woman. The proportion of preventive care manoeuvres carrying grade A and B recommendations (good or fair evidence for inclusion in the periodic health examination) from the Canadian Task Force on the Periodic Health Examination[13] that were performed, inquired about or offered ranged from 36% to 52%. The 48-year-old man had both more effective and more ineffective care than the other standardized patients. Ordering of tests increased with increasing age. Only the 28-year-old woman (who requested a prescription for birth control pills) was likely (in 92.7% of encounters) to get a prescription, although some physicians (14.8% of encounters) prescribed for the 70-year-old man.

Although the physicians usually used only one billing code, mean charges varied across patients. Within patient scenarios, the variation in physicians' charges ranged from nearly 6-fold to over 10-fold. The cost per minute of time spent was highest for the 28-year-old woman and lowest for the 70-year-old man. Two physicians reported detecting standardized patients; the rate of suspecting standardized patients was higher, usually because the patient did not fit the physician's practice profile in terms of such characteristics as age, sex, ethnicity and income group.

We examined the extent to which charges varied for a particular patient scenario using the following variables: time spent in office, preventive care performance score, diagnostic tests ordered (Yes), certification in family medi-

cine (Yes), female physician (Yes), group practice (Yes), graduated in 1980s (Yes), suspected standardized patient (Yes), fee-for-service remuneration (Yes) and problem-based medical school (Yes). Forward stepwise regression identified variables that accounted for a significant portion ($p \leq 0.05$) of the observed variability in charges. The results of the analysis are expressed as β coefficients, which are the amount of change in the dependent variables (cost in dollars, expressed in SD units) resulting from a change of 1 SD in each independent variable (e.g., time spent with the patient). Thus, the magnitude of the coefficient is directly interpretable as the relative strength of association. Time spent with the patient was a consistent predictor of physicians' charges (Table 2). Holding family medicine certification was associated with lower charges in 2 of the scenarios. Ordering diagnostic tests was associated with higher charges for the 48-year-old man. Between 33.2% to 55.5% of the variability in charges within patient scenarios could be explained.

Male physicians spent significantly more time with the 28-year-old woman (mean 14 [SD 10.6] minutes) than did female physicians (mean 10 [SD 6.3] minutes) ($p = 0.05$). No difference was seen in time spent with patients by sex of the physician for any of the other scenarios.

Repeated-measure ANCOVA showed that although some variability in charges was explained by individual physician differences ($p = 0.07$) and patient scenario ($p = 0.06$), other physician- and encounter-related factors were more significant. Time spent by the physician was the most significant factor explaining in-office service charges ($p < 0.01$), although the effect of time spent varied significantly across encounters, as evidenced by a significant time × standardized patient interaction ($p < 0.01$). Fee-for-service physicians charged more than physicians who usually had alternative billing arrangements ($p < 0.01$), and female physicians charged more than male physicians ($p = 0.03$). The effect of diagnostic testing on charges varied across patient scenarios, again shown by a significant interaction with standardized patient ($p = 0.01$). The remaining variables did not contribute significantly to explaining cost.

There were 59 billing codes for the 28-year-old woman, 66 for the 48-year-old man, 54 for the 52-year-old woman and 71 for the 70-year-old man. Table 3 presents the billing codes used to characterize encounters. General assessments (which can be used to bill for an annual health examination) were charged in a minority of cases overall but in more than half of the encounters with the 70-year-old man. A general assessment is the most expensive service, followed closely by counselling or psy-

Table 1: Characteristics of physicians' encounters with standardized patients, by patient scenario

Variable	Standardized patient scenario			
	28-year-old woman $n = 55$	48-year-old man $n = 54$	52-year-old woman $n = 52$	70-year-old man $n = 54$
Mean no. of minutes per encounter (and SD*)	12.3 (9.9)	22.3 (13.3)	15.8 (8.5)	24.6 (14.4)
Mean % of grade A and B manoeuvres† performed (and SD)	39.0 (18)	52.0 (19)	36.0 (16)	37.0 (15)
Mean % of grade D and E manoeuvres‡ performed (and SD)	7.0 (6)	37.0 (32)	9.0 (15)	17.0 (19)
% of encounters in which tests were ordered	22.0	35.2	34.6	44.4
% of encounters leading to a prescription	92.7	0.0	2.0	14.8
Mean no. of billing codes used (and SD)§	1 (0)	1 (0)	1 (0)	1 (0)
% of physicians who suspected or detected standardized patient	11.1	20.8	25.0	22.6
Mean amount billed (and SD), $§	34.40 (16.62)	40.55 (22.04)	34.64 (16.60)	48.20 (28.92)
Range of billing, $§	15.28–89.77	15.28–119.57	16.25–93.59	15.28–161.80
Mean cost/min (and SD), $	4.26 (3.40)	3.19 (6.79)	2.47 (1.11)	2.42 (2.07)

*SD = standard deviation.
†Good or fair evidence for inclusion in a periodic health examination.
‡Good or fair evidence for exclusion from a periodic health examination.
§Includes all in-office services (codes for minor, intermediate or general assessments, general reassessment, counselling, individual psychotherapy and in-office diagnostic services).

Table 2: Results of forward stepwise regression analysis to identify factors related to physicians' charges for the encounters, by patient scenario

Predictor variable	Standardized patient scenario							
	28-year-old woman		48-year-old man		52-year-old woman		70-year-old man	
	β	p value	β	p value	β	p value	β	p value
Time spent with patient	0.634	< 0.001	0.573	< 0.001	0.587	< 0.001	0.692	< 0.001
Certification in family medicine			−0.251	0.009			−0.310	0.002
Diagnostic tests ordered			0.347	0.006				
Adjusted R^2, %	39.1		55.5		33.2		53.0	

chotherapy for a half hour. The charges for intermediate assessments, done in about half of all encounters, are slightly more than half the charges for a general assessment. Minor assessments, which are supposed to include a brief assessment or brief advice or information, or both, cost about one-third of the charges for a general assessment. Minor assessments were charged most frequently for the 28-year-old woman.

More than two-thirds (67.5%) of the female physicians billed for counselling or psychotherapy, compared with 50.0% of the male physicians; however, the difference was not statistically significant. Across patients, female physicians were more likely than male physicians to charge for counselling or psychotherapy, but only for the 70-year-old man was this difference significant (27.3% v. 2.3%) (χ^2 = 8.69, p = 0.003).

Table 4 shows the diagnoses used to characterize the encounters. Annual health examination (no diagnosis) was used most often to characterize the 70-year-old man's encounters. (All other codes require a diagnosis.) The diagnostic codes assigned often fit with the standardized patients' stories (e.g., menopause-related diagnosis for the older woman, osteoarthritis for the older man). As well, diagnoses reflecting ill-defined conditions or signs and symptoms, including minor psychiatric problems, were assigned. Essential hypertension featured among the "other" diagnoses assigned to the 2 men.

Discussion

The Ontario Schedule of Benefits[14] does not provide a fee code for visits by new patients. Although most physicians in our study inquired about the personal and family history, performed some examination of the patient and provided some preventive health care, the variable time spent was reflected in their choice of billing code. About half of the physicians chose to characterize their services as intermediate assessments rather than the more expensive general assessment. The need to supply a diagnosis for every visit led to sometimes creative use of the diagnostic classifications available. Although charges correlated with time spent, there was a significant difference in charges for time spent across the 4 standardized patients. Physicians earned the most per minute for seeing the younger woman and least for seeing the older man. Furthermore, the time spent did not necessarily represent

Table 4: Commonly used diagnostic categories, by patient scenario

Patient scenario; diagnostic category	No. of physicians
28-year-old woman	
Annual health examination	2
Family planning/contraceptive advice	39
Signs/symptoms only (ill-defined condition)	5
Anxiety neurosis, hysteria, neurasthenia or reactive depression	6
Other	7
48-year-old man	
Annual health examination	8
Related to tobacco use	10
Signs/symptoms only (ill-defined condition)	11
Anxiety neurosis, hysteria, neurasthenia or reactive depression	10
Other (including 6 essential hypertension diagnoses)	27
52-year-old woman	
Annual health examination	4
Menopause-related diagnosis	28
Signs/symptoms only (ill-defined condition)	5
Anxiety neurosis, hysteria, neurasthenia or reactive depression	12
Other	4
70-year-old man	
Annual health examination	15
Related to osteoarthritis (knee/leg)	15
Signs/symptoms only (ill-defined condition)	7
Anxiety neurosis, hysteria, neurasthenia or reactive depression	10
Other (including 18 essential hypertension diagnoses)	24

Table 3: Frequently used billing codes (excluding diagnostic testing codes), by patient scenario

Billing code	Standardized patient scenario; no. (and %) of encounters			
	28-year-old woman $n = 55*$	48-year-old man $n = 54$	52-year-old woman $n = 52$	70-year-old man $n = 54$
General assessment	12 (22.2)	16 (29.6)	8 (15.4)	32 (59.2)
Counselling	9 (16.7)	13 (24.1)	15 (28.8)	5 (9.2)
Psychotherapy	1 (1.8)	2 (3.7)	1 (1.9)	2 (3.7)
Intermediate assessment	27 (50.0)	29 (53.7)	25 (48.1)	29 (53.7)
Minor assessment	9 (16.7)	6 (11.1)	5 (9.6)	3 (5.6)
General reassessment†	1 (1.8)	0 (0.0)	0 (0.0)	0 (0.0)
Second visits for physicians who billed	5 (9.2)	11 (20.4)	2 (3.8)	15 (27.8)
Total no. of visits	60	65	54	69

*One physician billed only for laboratory services; no consultation service was billed.
†Combined with intermediate assessment in analysis of billing codes used by sex of physician.

value for money. We noted no relation between the extent to which physicians followed the recommendations of the Canadian Task Force on the Periodic Health Examination[13] and their charges for the consultation.

Female physicians charged more for their services than their male colleagues, which suggests that they value their services more highly. They used the counselling/psychotherapy fee codes to characterize their services more often than did the male physicians. Other investigators who have found higher charges by female physicians have attributed this finding to different patient profiles and greater time spent with the patient.[15–18] In our study the female and male physicians spent similar amounts of time with patients. In fact, the male physicians spent significantly more time with the younger woman than did their female colleagues. We do not know whether our finding applies only to the cases seen (healthy new patients for whom inquiry and counselling regarding preventive care featured in many visits).

The charges of fee-for service physicians were higher than those of physicians who usually use alternative billing methods to charge the health insurance plan for services. This occurred despite our excluding from the analysis physicians who did not bill. Capitated physicians can charge fees for patients who are not included on their roster.

Our study has several limitations. The participants were a highly select group of volunteers. They participated in the earlier preventive care survey, had open practices and agreed to see unannounced standardized patients. The group's homogeneity may have limited our ability to detect differences. Furthermore, because there is no billing code for consultations with healthy new patients, this may not be the best standardized patient scenario to use to examine charges. However, when a patient presents with an illness, physicians also have discretion regarding their services. We did not include all visit costs in our analysis. Other charges (e.g., out-of-office diagnostic tests) and costs to the patient (e.g., prescription costs) were not included.

Our study shows the inherent weakness of fee schedules in billing for patient consultations. The billing codes allow for considerable physician latitude. Differences in cases related less to cost than did physician-associated characteristics. Charges varied widely across physicians and were not associated with quality of care. Unfortunately, alternative payment schemes also do not address variation in service provision among physicians. Only quality-assurance programs that provide guidelines for practice and monitor physician compliance are likely to increase value for money spent in health care. Publication of guidelines alone is ineffective, given the limited extent to which physicians adhered to the guidelines of the

Canadian Task Force on the Periodic Health Examination in the accompanying study (page 185). Further work is needed to find better approaches to improving practice and to ensure that the health care dollars spent represent value for money.

References

1. Vu NV, Barrows HS. Use of standardized patients in clinical assessments: recent developments and measurement findings. *Educ Res* 1994;23(3):23-30.

2. Van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med* 1990;2(2):58-76.

3. Owen A, Winkler R. General practitioners and psycho-social problems: an evaluation using pseudo patients. *Med J Aust* 1974;2:393-8.

4. Renaud M, Beauchemin J, Lalonde C, Poirier C, Berthiaume S. Practice settings and prescribing profiles: the simulation of tension headaches to general practitioners working in different practice settings in the Montreal area. *Am J Public Health* 1980;70:1068-73.

5. Woodward CA, McConvey CA, Neufeld V, Norman GR, Walsh A. Measurement of physician performance by standardized patients: refining techniques for undetected entry in physicians' offices. *Med Care* 1985;23:1019-27.

6. Norman GR, Neufeld VR, Walsh A, Woodward CA, McConvey GA. Measuring physicians' performance by using simulated patients. *J Med Educ* 1985; 60:925-34.

7. Russell NK, Boekeloo BO, Rafi IZ, Rabin DL. Using unannounced standardized patients to evaluate sexual risk assessment and risk reduction skills of practising physicians. *Acad Med* 1991;66(9, suppl):S37-9.

8. Hazelkorn HM, Macek MD. Perception of the need for removal of impacted third molars by general dentists and oral and maxillofacial surgeons. *J Oral Maxillofac Surg* 1994;52:681-6.

9. Hazelkorn HM. A comparison of dental treatment plans under different reimbursement systems. *J Public Health Policy* 1985;6:223-35.

10. Carney PA, Dietrich AJ, Freeman DH, Mott LA. The periodic health examination provided to asymptomatic older women: an assessment using standardized patients. *Ann Intern Med* 1993;119:129-35.

11. Black CD, Roos NP, Burchill CH. Utilization of hospital resources. *Med Care* 1995;33(suppl):DS55-72.

12. Wennberg J, Freeman J, Culp W. Are hospital services rationed in New Haven or over-utilized in Boston? *Lancet* 1987;1:1185-9.

13. Canadian Task Force on the Periodic Health Examination. *The Canadian guide to clinical preventive health care*. Ottawa: Health Canada; 1994. Cat no H21-117/1994E.

14. *Schedule of benefits: physician services under the Health Insurance Act*. Toronto: Ontario Ministry of Health; 1991.

15. Langwell KM: Factors affecting the incomes of men and women physicians: further explorations. *J Hum Resour* 1982;17:261-75.

16. Woodward CA, Hurley J. Comparison of activity level and service intensity of male and female physicians in five fields of medicine in Ontario. *Can Med Assoc J* 1995;153:1097-106.

17. Bertakis KD, Helms LJ, Callahan EJ, Azari R, Robbins JA. The influence of gender on physician practice style. *Med Care* 1995;33:407-16.

18. Keane D, Woodward CA, Ferrier BM, Cohen M, Goldsmith C. Female and male physicians. Different practice profiles. *Can Fam Physician* 1991;37:72-81.

Reprint requests to: *Dr. Christel A. Woodward, Rm. 3H4, Health Sciences Centre, McMaster University, 1200 Main St. W, Hamilton ON L8N 3Z5*