

Long-term outcomes for surgeons from 3- and 4-year medical school curricula

Jocelyn Lockyer, PhD^{*†}
 Claudio Violato, PhD^{*}
 Bruce Wright, MD[‡]
 Herta Fidler, MSc[†]
 Robert Chan, MD[§]

From the ^{*}Department of Community Health Sciences, [†]Continuing Medical Education and Professional Development, the [‡]Department of Family Medicine and Undergraduate Medical Education, University of Calgary, and [§]orthopedic surgeon, Edmonton, Alta.

Accepted for publication
 Mar. 24, 2011

Correspondence to:

J. Lockyer
 Senior Associate Dean, Education
 Faculty of Medicine
 University of Calgary
 3280 Hospital Dr. NW
 Calgary AB T2N 4Z6
 lockyer@ucalgary.ca

DOI: 10.1503/cjs.037610

Background: New approaches are needed to ensure that surgical trainees attain competence in a timely way. Traditional solutions have focused on the years spent in surgical training. We sought to examine the outcomes of graduates from 3-year versus 4-year medical schools for differences in surgeon performance based on multisource feedback data.

Methods: We used data from the College of Physicians and Surgeons of Alberta's Physician Achievement Review program to determine curricular outcomes. Data for each surgeon included assessments from 25 patients, 8 medical colleagues and 8 non-physician coworkers (e.g., nurses), and a self-assessment. We used these data to compare 72 physicians from a 3-year school matched with graduates from 4-year schools. The instruments were assessed for evidence of validity and reliability. We compared the groups using 1-way analysis of covariance and multivariate analysis of covariance, with years since graduation as a covariate, and a Cohen *d* effect size calculation to assess the magnitude of the change.

Results: Data for 216 surgeons indicated that there was evidence for instrument validity and reliability. No significant differences were found based on the length of the undergraduate program for any of the questionnaires or factors within the questionnaires.

Conclusion: Reconsideration might be given to the time spent in medical school before surgical training if training in the specialty and career years are to be maximized. This assumes that students are able to make informed career decisions based on clerkship and other experiences in a 3-year setting.

Contexte: Nous devons nous doter de nouvelles approches si nous voulons nous assurer que les résidents en chirurgie deviennent rapidement compétents. De tout temps, la solution a reposé sur le nombre d'années consacrées à la formation en chirurgie. Nous avons voulu comparer les résultats des diplômés d'une faculté de médecine dont le programme s'étale sur 3 ans à ceux des facultés de médecine dont le programme s'étale sur 4 ans pour vérifier les différences de rendement des chirurgiens à partir de données de sources diverses.

Méthodes: Nous avons utilisé les données du programme Physician Achievement Review du Collège des médecins et chirurgiens de l'Alberta pour mesurer les résultats. Les données pour chaque chirurgien incluait les évaluations de 25 patients, de 8 collègues médecins et de 8 collaborateurs non médecins (p. ex., infirmières), de même que les données d'une auto-évaluation. Nous avons utilisé ces données pour comparer 72 médecins d'une école dont le programme s'étale sur 3 ans assortis à des diplômés de facultés de médecine dont le programme s'étale sur 4 ans. Les instruments ont été évalués sur les plans de leur validité et de leur fiabilité. Nous avons comparé les groupes à l'aide d'analyses simples et multivariées de la covariance, avec le nombre d'années écoulées depuis l'obtention du diplôme comme covariable, et nous avons calculé la taille d'effet *d* de Cohen pour mesurer l'ampleur de la différence.

Résultats: Les données concernant 216 chirurgiens ont confirmé la validité et la fiabilité de l'instrument. On n'a observé aucune différence significative en lien avec la durée du programme de premier cycle pour aucun des questionnaires ou facteurs à l'intérieur des questionnaires.

Conclusion: Pour maximiser la spécialisation et les années de carrière, on pourrait revoir la durée de la formation en médecine menant à la spécialisation en chirurgie. Cela repose sur la prémisse selon laquelle les étudiants peuvent prendre des décisions éclairées quant à leur choix de carrière à partir de leur stage clinique et des autres expériences qu'ils acquièrent dans le contexte d'une formation étalée sur 3 ans.

Leaders and funders of medical education in both Canada and the United States are concerned about the length of time needed to graduate physicians from medical school and prepare them for competent independent medical and surgical practice.^{1,2} This is particularly problematic for surgery, which has a long period of training after medical school. Surgical trainees are increasingly expected to extend their core training through subspecialization,^{3,4} and there is concern that reduced work hours in training programs will compromise the attainment of surgical skills.^{5,6}

The Royal College of Physicians and Surgeons of Canada (RCPSC)⁷ requires that attention be paid to core competencies in addition to surgical skills. Addressing these competencies has been challenging for some programs⁸ and requires creative approaches. These approaches have included simulation⁹ and assessment with the objective structured clinical examination,¹⁰ and a move away from time-based training to competency-based training.¹¹ Time spent in residency is precious and must be used to its maximum advantage.¹²

Solutions to surgical competency have traditionally addressed the years spent in surgical training. Little attention has been paid to the time before surgical residency, namely medical school, and whether it can be shortened to maximize the time available for development of surgical skills. During the 1970s, almost one-quarter of the medical schools in the United States established 3-year medical school programs in response to federal legislation providing financial incentives.² These programs were later discontinued, despite the lack of evidence that students were at a disadvantage.² By contrast, Canada established 3-year programs at both the University of Calgary and McMaster University. The University of Calgary curriculum is 131 weeks compared with the University of Alberta curriculum of 144 weeks. At the University of Calgary, the first 2 years combine basic and clinical science in a clinical presentation curriculum, and the final year is a traditional clinical clerkship. At the University of Alberta, the first 2 years focus on organ systems, the third year provides rotations through the standard 6 specialties and the fourth-year rotations provide a deeper understanding of subspecialties.¹² Other Canadian 4-year medical schools would have curricula similar to that of the University of Alberta. This 30-year natural experiment with graduates of 3-year medical programs appears to have produced physicians who are competent,^{1,12,13} although the objective data come from performance data for family physician-general practitioners^{12,13} and not surgeons.

The purpose of this study was to compare the performance of practising surgeons in Alberta who graduated from the University of Calgary (a 3-year school) with matched samples from the University of Alberta (a 4-year school) and graduates from other 4-year Canadian med-

ical schools. We used data from a regulatory authority that assesses a broad range of competencies related to clinical skill, communication, professionalism and management skill. In Alberta, each surgeon must participate in the Physician Achievement Review (PAR) program, a province-wide multisource evaluation performed every 5 years. This evaluation is required by the College of Physicians and Surgeons of Alberta (CPSA), the regulatory licensing authority, and consists of questionnaire data from 25 patients, 8 medical colleagues and 8 non-physician coworkers (e.g., nurses, physical therapists), as well as self-assessment data. All of the surgical subspecialties are assessed using 1 set of surveys. Participation in PAR every 5 years along with participation in the maintenance of certification program of the RCPSC is required for continued licensure. When the questionnaires were developed, almost 10 years ago, they showed evidence for validity and reliability.^{14,15} Since the program became mandatory, 645 surgeons have participated in the program at least once, 383 of whom have participated on 2 occasions.

This study addressed the following questions.

1. What is the current evidence for the reliability and validity of the PAR questionnaires used for the surgical specialties based on data from the surgeons' most recent assessment?
2. What are the factors within each of the questionnaires used for the surgical specialists?
3. Are there differences in mean aggregate scores between schools (University of Calgary, University of Alberta and other Canadian 4-year medical schools) based on the whole assessment and on the factors within questionnaires?

METHODS

Pivotal Research Inc., the administrator of the CPSA's PAR program, under the direction of the CPSA, provided an anonymous data set for this study. The data set consisted of a sample of 550 Canadian surgeons who completed the PAR program. Each University of Calgary graduate in the data set ($n = 72$) was matched to graduates from the same year (or close to the same year) who graduated from the University of Alberta and other Canadian medical schools. Where there was more than 1 possible match, the match was made to someone in the same subspecialty (e.g., ophthalmology). Graduates of McMaster University were excluded as they have a 3-year curriculum. Graduates of international medical schools were excluded because their undergraduate education is likely to be more variable than that of graduates of schools accredited by the Liaison Committee on Medical Education.

Data for each physician included assessments from 25 patients, 8 medical colleagues and 8 nonphysician

coworkers, and a self-assessment. Whereas the questionnaires were not based on CanMEDS competencies, the items reflect many of the CanMEDS roles, namely Medical Expert, Scholar, Professional, Communicator, Collaborator, Manager and Health Advocate. They were items that could be observed and captured the public and professional expectations of a surgeon. Copies of the questionnaires are available online (www.par-program.org). The assessments use a 5-point Likert scale. All questionnaire forms provide the respondent with an “unable to assess” option. The data set included a limited amount of sociodemographic information (year and school of graduation, urban/regional/rural location, specialty). The data were from surgeons’ most recent PAR experience (collected between 2002 and 2009).

The data were analyzed in a number of ways. First, because the instruments were developed and assessed almost 10 years ago, a current psychometric assessment was deemed appropriate to ensure the instruments and their items continued to provide evidence of validity and reliability. This is important for this study and for the future use of the instruments. Other regulatory jurisdictions have adopted or are considering using the instruments; it is important that they be assessed regularly. Descriptive data calculations were done for each of the items on each of the questionnaires. These analyses enabled an examination of the range and mean (and standard deviation [SD]) for each item. As well, items that are not functioning well (i.e., have high percentage of “unable to assess”) were identified. A reliability analysis was completed by calculating the Cronbach α for each survey and each factor to determine the internal consistency reliability of the instruments and scales. A generalizability study (E_p^2) for each survey was conducted to establish the reliability of the data for each surgeon who was assessed. This assessment indicates whether the combination of items and raters achieves an appropriate level (generally $E_p^2 = 0.70$). These data informed question 1. A confirmatory factor analysis for each of the questionnaires determined the factor structures for each instrument (research question 2). Last, a 1-way analysis of covariance (ANCOVA) was used to evaluate differences between schools for each instrument using the aggregate mean questionnaire score as the dependent variable and years since graduation as the covariate (to control for potential confounding of effect by years since graduation). A multivariate analysis of covariance (MANCOVA) was used to evaluate the differences between medical schools on the aggregate mean factor scores for each questionnaire, with years since graduation as the covariate. The significance level was calculated using multivariate F from the ANCOVA and MANCOVA with a significance level set at 0.05. An effect size calculation (Cohen d) was used to determine the magnitude of differences (research question 3).

RESULTS

Participant matching produced a data set of 216 surgeons (72×3). As shown in Table 1, the groups were about equal in terms of response rates on questionnaires. There were more men in the University of Alberta cohort (86.1%) than in the cohorts from the University of Calgary (80.6%) and other medical schools (73.6%). Similarly, whereas the numbers of surgeons practising in urban centres were similar, there were more University of Alberta graduates practising in regional centres than from the other 2 groups.

The medical colleague data indicated that the mean score for all items was greater than 4.0 out of 5. There were a few items that a higher percentage of colleagues were unable to assess. These were mostly items that were difficult to observe easily (e.g., medical record quality, professional development involvement, stress management, areas beyond scope, critical evaluation of medical literature and contribution to quality improvement). The self-assessment questionnaire, written in the first person, had items that were identical to the medical colleague assessment. The self-assessment scores the surgeons provided were lower on all items, with means ranging from 3.58 for the item on contribution to quality improvement activities to 4.30 for respecting the rights of patients. The Cronbach α scores for both the medical colleague and self-assessments were high, at greater than 0.9. The generalizability coefficient study resulted in a coefficient of $E_p^2 = 0.61$. A 4-factor solution emerged from the analysis, which explained almost 75% of the variance. The items aligned into 4 broad areas (factors): communication and professionalism, medical expert, scholar and manager. These data are provided in Table 2.

Table 1. Sociodemographic comparison of graduates of the University of Calgary, University of Alberta and other Canadian universities

Assessment	Calgary, <i>n</i> = 72	Alberta, <i>n</i> = 72	Other, <i>n</i> = 72
No. of self-assessments	72	72	72
Medical colleagues			
No. of responses	554	552	551
Response rate per physician	7.70	7.67	7.65
Nonphysician coworkers			
No. of responses	550	551	538
Response rate per physician	7.64	7.70	7.47
Patients			
No. of responses	1733	1719	1731
Responses rate per physician	24.06	23.88	24.04
Mean years of practice at time of review	19.21	19.24	19.15
Percent male	80.6	86.1	73.6
Location of practice, %			
Urban	79.2	77.8	81.9
Rural	8.3	2.8	6.9
Regional	12.5	19.4	11.1

Table 2. Descriptive statistics and factor analysis for assessment by medical colleagues and self-assessment

Colleague item	Medical colleagues			Self			Factor‡			
	No. of surgeons*	Mean (SD) score†	%UA	No. of surgeons*	Mean (SD) score†	%UA	1	2	3	4
1. Communicates effectively with patients	216	4.55 (0.341)	2.4	215	4.29 (0.692)		0.79	0.21	0.13	0.30
2. Communicates effectively with patients' families	216	4.52 (0.373)	12.3	216	4.10 (0.695)		0.77	0.17	0.18	0.34
3. Communicates effectively with other health care professionals	216	4.57 (0.343)	0.4	216	4.04 (0.734)		0.80	0.27	0.18	0.14
4. Within the range of services provided by this physician, he/she performs technical procedures skillfully	216	4.71 (0.286)	1.8	215	4.25 (0.761)		0.19	0.81	0.27	-0.03
5. Selects diagnostic tests appropriately	216	4.61 (0.242)	6.8	215	4.00 (0.758)	0.5	0.19	0.77	0.27	0.27
6. Critically assesses diagnostic information	216	4.63 (0.242)	3.6	216	4.02 (0.744)		0.24	0.75	0.31	0.23
7. Makes the correct diagnosis in a timely fashion	215	4.67 (0.224)	2.4	216	4.13 (0.740)		0.27	0.78	0.24	0.23
8. In general, selects appropriate treatments	216	4.68 (0.248)	1.9	215	4.20 (0.750)		0.36	0.71	0.30	0.20
9. Maintains quality medical records	216	4.49 (0.322)	21.1	216	3.81 (0.861)		0.13	0.13	0.21	0.78
10. Handles transfer of care appropriately	216	4.53 (0.300)	10.3	213	3.91 (0.750)	1.4	0.60	0.29	0.14	0.53
11. Provides a clear understanding about who is responsible for continuing care of the patient	216	4.55 (0.286)	6.7	216	4.00 (0.741)	0.2	0.48	0.38	0.15	0.56
12. Communicates information to patients about rationale of treatment	216	4.58 (0.286)	7.2	215	4.28 (0.689)	0.5	0.64	0.33	0.23	0.41
13. Recognizes psychological aspects of illness	216	4.35 (0.348)	16.1	216	3.87 (0.775)		0.68	0.00	0.21	0.36
14. Maintains confidentiality of patients and their families	216	4.60 (0.246)	12.4	215	4.27 (0.706)	0.5	0.51	0.42	0.23	0.30
15. Coordinates care effectively for patients with other health care professionals	216	4.58 (0.271)	4.0	214	3.88 (0.747)	0.9	0.66	0.38	0.23	0.37
16. Coordinates the management of care for patients with complex problems	215	4.58 (0.281)	8.0	205	3.89 (0.706)	5.1	0.50	0.42	0.40	0.34
17. Respects the rights of patients	216	4.59 (0.254)	4.6	216	4.30 (0.713)		0.69	0.40	0.16	0.25
18. Collaborates with medical colleagues	216	4.59 (0.297)	0.9	213	4.08 (0.723)	1.4	0.75	0.27	0.30	0.11
19. Is involved with professional development	215	4.55 (0.303)	16.8	213	4.02 (0.783)	1.4	0.25	0.20	0.82	0.15
20. Accepts responsibility for own professional actions	216	4.62 (0.252)	3.7	215	4.29 (0.704)	0.5	0.60	0.50	0.41	0.13
21. Manages health care resources efficiently	216	4.47 (0.301)	13.1	216	3.88 (0.760)		0.49	0.53	0.22	0.31
22. Manages stress effectively	216	4.33 (0.421)	18.9	216	3.62 (0.860)		0.62	0.41	0.18	-0.04
23. Participates in a system of call to provide care for his/her patients when unavailable	216	4.60 (0.287)	8.6	207	4.19 (0.743)	3.7	0.49	0.51	0.11	0.25
24. Recognizes his/her own surgical limitations	216	4.60 (0.258)	7.4	216	4.19 (0.711)		0.47	0.67	0.20	0.23
25. Handles requests for consultation in a timely manner	216	4.51 (0.284)	5.7	215	3.80 (0.820)	0.5	0.36	0.50	0.11	0.45
26. Advises referring physician if referral request is outside of scope of practice	216	4.58 (0.282)	19.1	213	3.98 (0.727)	1.4	0.26	0.56	0.14	0.53
27. Assumes appropriate responsibility for patients	216	4.63 (0.253)	1.0	215	4.20 (0.714)	0.5	0.59	0.48	0.29	0.35
28. Provides timely information to referring physicians about mutual patients	216	4.57 (0.279)	9.9	216	3.85 (0.783)		0.41	0.46	0.23	0.57
29. Critically evaluates the medical literature to optimize clinical decision making	215	4.53 (0.323)	19.1	213	3.82 (0.805)	1.4	0.12	0.34	0.76	0.27
30. Facilitates the learning of medical colleagues and coworkers	216	4.53 (0.330)	9.9	212	3.88 (0.763)	1.9	0.36	0.31	0.72	0.07
31. Contributes to quality improvement programs and practice guidelines	215	4.49 (0.368)	27.3	201	3.58 (0.951)	6.9	0.30	0.30	0.74	0.19
32. Participates effectively as a member of the health care team	216	4.59 (0.311)	1.0	212	4.00 (0.729)	1.4	0.78	0.34	0.32	0.08
33. Exhibits professional and ethical behaviour toward physician colleagues	216	4.62 (0.318)	0.1	214	4.13 (0.749)	0.9	0.84	0.29	0.21	0.03
34. Shows compassion for patients and their families	216	4.58 (0.317)	4.2	216	4.28 (0.721)		0.86	0.20	0.13	0.20
% variance explained (total = 73.246)							58.636	6.746	4.263	3.602
Cronbach α	Colleague, all items = 0.980			Self, all items = 0.975			0.969	0.937	0.894	0.801

SD = standard deviation; %UA = percent unable to assess.
 *The numbers of surgeons differ because some respondents were unable to assess some items.
 †On a 5-point Likert scale.
 ‡Factor 1 = communication and professionalism; factor 2 = medical expert, factor 3 = scholar, factor 4 = manager. The use of bold for factor loading indicates the best loading.

The nonphysician coworkers gave scores between 4.36 and 4.66 on all items. Three items had high “unable to assess” rates and were related to written information about prescriptions and hospital orders as well as accessibility for communication about mutual patients. The internal consistency reliability was 0.955 for the whole scale. The generalizability coefficient was $Ep^2 = 0.70$. A 2-factor solution (oral communication and professionalism, and written communication) emerged, which explained 72% of the variance. These data are provided in Table 3.

The means on the patient data ranged from 4.34 to 4.78. Patients were able to assess most items. The internal consistency reliability for the overall questionnaire was greater than 0.98. The generalizability study resulted in a coefficient of $Ep^2 = 0.81$. A 4-factor solution emerged from the analysis, accounting for 77% of the variance. The 4 factors were communication, manager, follow-up and management. These data are provided in Table 4.

The comparison of the aggregate mean scores and mean factor scores showed that there were no differences by school for any of the assessments or factors within the questionnaires (Table 5).

DISCUSSION

This study affirms that the multisource feedback instruments developed for the PAR program to assess surgeons

are still viable after nearly 10 years of use. There are relatively few items that medical colleagues, nonphysician coworkers and patients are unable to assess. The items that have the most “unable to assess” responses are ones that are difficult to observe and likely should be revised or deleted. The instruments and their scales are reliable, as shown by the Cronbach α analysis. The generalizability study indicates that all 3 instruments reached stability, but the medical colleague instrument was less reliable than the nonphysician coworker and patient instruments.

The factor analysis indicates that the items continue to correlate in ways that reflect CanMEDS roles. Surgeons participating in the program receive feedback about their oral and written communication from medical colleagues, nonphysician coworkers and patients. They also receive feedback to varying degrees about the Medical Expert and Scholar roles from medical colleagues, manager skills from medical colleagues and patients, and professionalism from medical colleagues and nonphysician coworkers.

The instruments have been adopted for use in Nova Scotia and are being considered for use in other Canadian jurisdictions. In Canada, regulatory authorities are increasing their expectations of physicians. Multisource feedback is a relatively inexpensive assessment method designed to provide physicians and surgeons with formative feedback to guide development in areas that would not be available to them through institutional surgical audits. In the United

Table 3. Descriptive statistics and factor analysis for assessment by nonphysician coworkers

Coworker item	No. of surgeons*	Mean (SD) score†	%UA	Factor‡	
				1	2
1. Communicates effectively with patients	216	4.59 (0.373)	3.5	0.70	0.46
2. Verbal communication with other health professionals is effective	216	4.50 (0.419)	1.3	0.87	0.22
3. Written communication with other health professionals is effective	216	4.46 (0.353)	15.3	0.58	0.60
4. Writes prescriptions clearly	215	4.37 (0.468)	22.9	0.14	0.89
5. Is courteous to coworkers	216	4.54 (0.497)	0.6	0.87	0.09
6. Respects the professional knowledge and skills of coworkers	216	4.56 (0.412)	1.0	0.87	0.18
7. Writes hospital orders clearly	215	4.39 (0.486)	28.1	0.13	0.89
8. Shows compassion to patients and their families	216	4.58 (0.359)	3.8	0.69	0.44
9. Is able to separate personal values from management of patients	216	4.52 (0.337)	13.2	0.74	0.44
10. Is courteous to patients and their families	216	4.64 (0.342)	2.4	0.79	0.39
11. Respects the rights of patients to make informed decisions	216	4.63 (0.318)	8.1	0.70	0.50
12. Accepts responsibility for professional actions	216	4.63 (0.302)	6.0	0.71	0.49
13. Accepts responsibility for patient care	216	4.66 (0.285)	2.6	0.66	0.54
14. Collaborates well with coworkers	216	4.47 (0.439)	1.2	0.91	0.21
15. Is reasonably accessible to patients	216	4.36 (0.373)	11.2	0.63	0.43
16. Maintains confidentiality of patients	216	4.69 (0.260)	6.6	0.41	0.55
17. Accessible for appropriate communication about mutual patients	215	4.53 (0.357)	21.2	0.69	0.40
18. Effectively participates as a member of a health care team	216	4.58 (0.362)	2.4	0.84	0.31
19. Facilitates the learning of medical colleagues and coworkers	216	4.58 (0.356)	8.3	0.69	0.23
% Variance explained (total = 71.879)				63.882	7.997
Cronbach α		All items = 0.955		0.968	0.834

SD = standard deviation; %UA = percent unable to assess.

*The numbers of surgeons differ because some respondents were unable to assess some items.

†On a 5-point Likert scale.

‡Factor 1 = oral communication and professionalism; factor 2 = written communication. The use of bold for factor loading indicates the best loading.

Kingdom, the National Health Service is working toward the inclusion of multisource feedback, using questionnaire data from colleagues and patients as part of its revalidation procedures in all specialties, including surgery.^{16,17} An instrument such as the one developed in Alberta may be suitable for an international environment. It may also be a

helpful way of enhancing behaviours related to professionalism, communication and collaboration.

We did not find any differences in surgeon performance by school of graduation either in the assessment as a whole or in factors within the questionnaires. This suggests an equivalency of performance for graduates of the University

Table 4. Descriptive statistics and factor analysis for assessment by patients

Patient item	No. of surgeons	Mean (SD) score*	%UA	Factor†			
				1	2	3	4
1. My surgeon explained my condition to me satisfactorily	216	4.71 (0.174)	3.2	0.79	0.29	0.23	0.19
2. Before booking my surgery, my surgeon explained my procedure thoroughly in language I understood	216	4.71 (0.182)	16.1	0.74	0.15	0.28	0.19
3. Before booking my surgery, my surgeon explained any alternatives thoroughly in language I understood	216	4.64 (0.194)	18.4	0.73	0.08	0.35	0.14
4. My surgeon explained what could be done if my illness was to recur	216	4.48 (0.264)	25.4	0.54	0.13	0.57	0.11
5. My surgeon or his/her staff explained when to return for follow-up care	216	4.68 (0.186)	9.4	0.59	0.42	0.37	0.16
6. My surgeon or his/her staff provided me with instructions on how and when to take my medicine	216	4.63 (0.233)	26.4	0.57	0.35	0.43	0.12
7. My surgeon told me of side effects of the medicine	216	4.58 (0.229)	17.6	0.70	0.24	0.40	0.10
8. Spends enough time with me	216	4.60 (0.205)	1.0	0.82	0.29	0.23	0.19
9. Shows interest in my problems	216	4.67 (0.190)	1.1	0.85	0.24	0.20	0.19
10. Asks appropriate details about my personal history	216	4.60 (0.197)	1.8	0.79	0.28	0.28	0.20
11. Answers my questions well	216	4.68 (0.183)	0.8	0.86	0.27	0.21	0.17
12. Examines me appropriately for my problems	216	4.70 (0.179)	1.9	0.82	0.34	0.21	0.20
13. Treats me with respect	216	4.76 (0.154)	0.8	0.83	0.34	0.15	0.20
14. Talks with me about treatment plan and alternatives	216	4.65 (0.187)	4.8	0.82	0.20	0.32	0.16
15. Is easy to get to (e.g., parking, wheelchair, etc.)	216	4.29 (0.358)	1.7	0.14	0.13	0.12	0.76
16. Has sufficient waiting areas	216	4.46 (0.346)	1.3	0.11	0.13	0.07	0.90
17. Examining rooms are adequately sized and have adequate equipment	216	4.58 (0.213)	1.7	0.31	0.35	0.25	0.76
18. Is clean and in good repair	216	4.62 (0.230)	0.8	0.20	0.32	0.17	0.82
19. Provides adequate privacy	216	4.59 (0.202)	1.2	0.32	0.38	0.18	0.73
20. I can reach the office by phone during the day	216	4.46 (0.282)	4.8	0.17	0.57	0.26	0.33
21. In an emergency situation, my surgeon's office provides me with clear instructions on what I am to do	216	4.44 (0.256)	9.6	0.36	0.53	0.53	0.15
22. My messages are returned	216	4.50 (0.258)	18.9	0.22	0.63	0.48	0.16
23. The staff is very capable	216	4.64 (0.196)	1.6	0.35	0.82	0.27	0.22
24. The staff is helpful and pleasant	216	4.67 (0.191)	0.5	0.29	0.84	0.22	0.25
25. The staff is respectful of patients	216	4.68 (0.182)	0.8	0.35	0.83	0.25	0.25
26. The staff behaves in a professional manner	216	4.68 (0.187)	0.8	0.34	0.83	0.25	0.26
27. The staff works well with my surgeon	216	4.66 (0.196)	9.5	0.37	0.75	0.32	0.23
28. Prevents patients from hearing confidential information about other patients	216	4.54 (0.214)	9.2	0.37	0.57	0.30	0.34
29. I receive an appropriate explanation if my appointment is delayed	216	4.34 (0.303)	29.8	0.20	0.45	0.50	0.31
30. When asked, my surgeon provides insurance and medicolegal reports	216	4.41 (0.306)	50.5	0.29	0.21	0.76	0.18
31. When asked, my surgeon provides reports, files, or copies of letters	216	4.48 (0.266)	45.0	0.26	0.31	0.76	0.22
32. I am advised of results of tests or x-rays	216	4.59 (0.206)	18.2	0.52	0.41	0.47	0.15
33. My surgeon arranges appointments with other specialists when necessary	216	4.55 (0.243)	40.1	0.46	0.26	0.65	0.11
34. Someone from my surgeon's office follows up on any serious problems I may have	216	4.49 (0.279)	39.1	0.31	0.36	0.75	0.12
35. I am told what to do if my problems do not get better	216	4.54 (0.232)	20.8	0.51	0.40	0.57	0.21
36. My doctor asks regularly about prescription and nonprescription medicine I may be taking	216	4.59 (0.198)	6.5	0.57	0.51	0.15	0.16
37. My doctor has printed health information available	216	4.53 (0.208)	17.4	0.52	0.51	0.30	0.25
38. I would go back to this doctor	216	4.78 (0.155)	1.7	0.68	0.55	0.13	0.15
39. I would send a friend to this doctor	216	4.77 (0.156)	1.8	0.69	0.52	0.12	0.16
% variance explained (total = 76.661)				60.172	7.590	5.008	3.891
Cronbach α		(All items = 0.987)		0.976	0.955	0.933	0.900

SD = standard deviation; %UA = percent unable to assess.
 *On a 5-point Likert scale.
 †Factor 1 = communication; factor 2 = manager; factor 3 = follow-up; factor 4 = management. The use of bold for factor loading indicates the best loading.

of Calgary and those from 4-year medical schools. Detecting differences is difficult, as the clerkship experience, residency program, postfellowship training, practice norms, patient expectations and changing science will all have an impact on differences over time.

Limitations

There are limitations to our study. Our study size of 72 in each group was small. We were limited by the numbers of University of Calgary graduates who trained in surgery and entered practice in Alberta. Those 72 surgeons provided the base for matching with University of Alberta graduates and those of other Canadian medical schools. There were insufficient numbers of McMaster graduates (the other 3-year medical school) practising as surgeons in Alberta to add them to the comparison. The ideal of randomization of the effect of curriculum on performance would be a preferable design to our covariate analysis, which used years since graduation as the covariate, but this design is not easily attainable using data from a naturalistic environment. Nonetheless, we tried to match the 3 groups of physicians and subsequently to covary the potential impact of years of practice on the dependent variables. The physicians in the study had been in practice for a mean of 19 years. We must also note that curricula of medical schools evolve, and making comparisons can be difficult.

CONCLUSION

The present study is unique. Once physicians get beyond certification and board examinations, few data exist to make long-term comparisons. These data provide a comparative quantitative analysis showing that graduating competent physicians in a shorter time is possible, and the shorter time does not appear to have long-term implications for surgical practice based on multisource feedback data for surgeons. Response rates from each source for each group were comparable and were high because participation in the PAR program is mandatory. The study shows that if the duration of medical school programs could be adjusted, graduates could train in surgery and be ready for practice 1 year earlier without any detectable differences in competency on our assessment tool. However, this assumes that students are able to make informed career decisions sufficiently early to plan clerkship electives and effectively assess their career choices.

Acknowledgements: The authors thank the College of Physicians and Surgeons of Alberta, particularly John Swiniarski, Trevor Theman and Bryan Ward (deceased) who enabled this study to be conducted, as well as Steven Dennis from Pivotal Research Inc.

Competing interests: None declared.

Contributors: J. Lockyer, C. Violato and B. Wright designed the study. J. Lockyer acquired the data, which was analyzed by J. Lockyer, C. Violato, H. Fidler and R. Chan. J. Lockyer, C. Violato and H. Fidler wrote the

Table 5. Aggregate mean item and factor scores by medical school of graduation

Item	Range of scores	Medical school; mean (SD) score			$F_{2,213}^*$	p value
		University of Calgary	University of Alberta	Other university		
Self aggregate mean item scores	1.73–5.00	4.04 (0.583)	4.07 (0.555)	3.99 (0.545)	0.345	0.71
Self aggregate mean factor scores						
Factor 1: Communication and professionalism	2.06–5.00	4.14 (0.583)	4.10 (0.544)	4.03 (0.557)	0.631	0.53
Factor 2: Medical expert	1.50–5.00	4.03 (0.647)	4.13 (0.606)	4.03 (0.590)	0.557	0.57
Factor 3: Scholar	1.50–5.00	3.81 (0.722)	3.81 (0.667)	3.85 (0.715)	0.062†	0.94
Factor 4: Manager	1.00–5.00	3.87 (0.739)	3.98 (0.651)	3.81 (0.625)	0.090	0.34
Medical colleague aggregate mean item scores	3.65–4.96	4.59 (0.196)	4.57 (0.205)	4.54 (0.265)	0.927	0.40
Medical colleague aggregate mean factor scores						
Factor 1: Communication and professionalism	3.55–4.97	4.60 (0.211)	4.55 (0.226)	4.51 (0.306)	2.152	0.12
Factor 2: Medical expert	3.69–4.99	4.61 (0.204)	4.62 (0.195)	4.58 (0.243)	0.764	0.47
Factor 3: Scholar	3.58–5.00	4.51 (0.273)	4.50 (0.315)	4.56 (0.281)	0.776	0.46
Factor 4: Manager	3.69–5.00	4.56 (0.225)	4.53 (0.233)	4.52 (0.288)	0.473	0.62
Nonphysician coworker aggregate mean item scores	3.63–4.98	4.60 (0.255)	4.52 (0.312)	4.51 (0.310)	2.088	0.13
Nonphysician coworker aggregate mean factor scores						
Factor 1: Oral communication and professionalism	3.32–4.97	4.62 (0.267)	4.54 (0.320)	4.52 (0.333)	1.938	0.15
Factor 2: Written communication	3.11–5.00	4.54 (0.304)	4.44 (0.353)	4.47 (0.321)	1.708	0.18
Patient aggregate mean item scores	4.08–4.92	4.60 (0.163)	4.57 (0.169)	4.60 (0.168)	0.989	0.38
Patient aggregate mean factor scores						
Factor 1: Communication	4.16–4.96	4.68 (0.147)	4.63 (0.171)	4.67 (0.164)	1.400	0.25
Factor 2: Manager	4.01–4.96	4.58 (0.198)	4.56 (0.191)	4.61 (0.179)	1.209	0.30
Factor 3: Follow-up	3.55–4.96	4.48 (0.223)	4.44 (0.235)	4.49 (0.217)	0.945	0.39
Factor 4: Management	3.00–4.93	4.49 (0.261)	4.51 (0.230)	4.52 (0.215)	0.256	0.77

SD = standard deviation.
 *Unless otherwise indicated.
 † $F_{2,212}$.

article, which was reviewed by J. Lockyer, C. Violato, B. Wright and R. Chan. All authors approved the article that was submitted for publication.

References

1. Flegel KM, Hébert PC, MacDonald N. Is it time for another medical curriculum revolution? *CMAJ* 2008;178:11.
2. Whitcomb ME. Who will study medicine in the future? *Acad Med* 2006;81:205-6.
3. Hudkins JR, Helmer SD, Smith RS. General surgery resident practice plans: A workforce for the future? *Am J Surg* 2009;198:798-803.
4. Richardson JD. Training of general surgical residents: What model is appropriate? *Am J Surg* 2006;191:296-300.
5. Payette M, Chatterjee A, Weeks WB. Cost and workforce implications of subjecting all physicians to aviation industry work-hour restrictions. *Am J Surg* 2009;197:820-5.
6. Griner D, Menon RP, Kowall CA, et al. The eighty-hour workweek: surgical attendings' perspective. *J Surg Educ* 2010;67:25-31.
7. Royal College of Physicians and Surgeons of Canada. *The CanMEDS 2005 Physician Competency Framework*. Ottawa (ON): The College; 2005. Available: www.rcpsc.medical.org/canmeds/CanMEDS2005/index.php (accessed 2010 May 18).
8. Yaszay B, Kubiak E, Agel J, et al. ACGME core competencies: Where are we? *Orthopedics* 2009;32:171-6.
9. Dutta S, Krummel M. Simulation: a new frontier in surgical education. *Adv Surg* 2006;40:249-63.
10. Hochberg MS, Kalet A, Zabar S, et al. Can professionalism be taught? Encouraging evidence. *Am J Surg* 2010;199:86-93.
11. Wade VM, Dev P, Buckley R, et al. Competencies for a Canadian orthopaedic surgery core curriculum. *J Bone Joint Surg Br* 2009;91:1618-22.
12. Lockyer JM, Violato C, Wright BJ, et al. An analysis of long-term outcomes of the impact of curriculum: a comparison of the three- and four-year medical school curricula. *Acad Med* 2009;84:1342-7.
13. Norman GR, Wenghofer E, Klass D. Predicting doctor performance outcomes of curriculum interventions: problem-based learning and continuing competence. *Med Educ* 2008;42:794-9.
14. Violato C, Lockyer JM, Fidler H. Multisource feedback: a method of assessing surgical practice. *BMJ* 2003;326:546-8.
15. Lockyer J, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. *Teach Learn Med* 2003;15:168-74.
16. General Medical Council. Annex 3 – GMC Principles, criteria and key indicators for colleague and patient questionnaires in revalidation. London (UK): The Council; 2010. In: *Revalidation: the way ahead*. Available: www.gmc-uk.org/static/documents/content/Revalidation_way_ahead_annex3.pdf (accessed 2012 May 12).
17. Royal College of Surgeons. *An overview of revalidation for surgeons*. London (UK): The College; 2010. Available: www.rcseng.ac.uk/revalidation/documents/overview_of_revalidation%20final.pdf (accessed 2012 May 12).



Call for papers: CMAJ Holiday Reading

Hilarity and good humour ... help enormously in both the study and the practice of medicine. — William Osler

Submit your twisted research, biting satire or culturally sensitive missives to CMAJ's award-winning, intensely cited Holiday Reading section. Articles should be no longer than 1200 words; photographs and illustrations are most welcome. Submit your article online at <http://mc.manuscriptcentral.com/cmaj>. If you need more information, kindly contact barbara.sibbald@cmaj.ca

Deadline: October 1, 2012