

# Evaluation of Residency Programs: A Novel Approach Using Simulation

KENNETH DOYLE, MD  
MEREDITH YOUNG, PhD  
SARKIS METERISSIAN, MDCM

## Abstract

**Background** In Canada, there has been discussion about restructuring the accreditation process for residency education programs with the possibility of exempting selected programs from regular, on-site, external reviews.

**Objective** We assessed the feasibility and acceptability of a structured and rigorous internal review that identified program strengths and weaknesses, with the aim of allowing well-performing programs to be exempt from external reviews or to facilitate a significant lengthening of the review cycle.

**Methods** We simulated all aspects of a regular, on-site, external review. All participants (program directors, program coordinators, faculty surveyors, and resident surveyors) were trained and performed all components of a formal external review. Participants completed an online survey to assess perceptions of the process and outcome.

**Results** The overall response rate was 73% (109 of 149). Most respondents perceived the process to be either extremely or very rigorous (84%), fair (82%), and unbiased (75%). Those with previous review experience (77%) reported that the internal review process simulated a regular, on-site, external review either well or very well (mean rating 4.87, SD 0.90). Most program directors reported the cited list of program strengths to be either extremely or very appropriate (74%, 26 of 35). Perceptions of fairness, bias, and the appropriateness of cited program strengths and weaknesses depended on review outcome.

**Conclusions** A structured and rigorous internal review process that simulates a regular, on-site, external review is feasible and could yield a list of program strengths and weaknesses for use in ongoing assessment and improvement.

*Editor's Note: The online version of this article contains the survey instruments used in this study.*

## Introduction

In Canada, residency training programs are accredited through the Royal College of Physicians and Surgeons of Canada, the College of Family Physicians of Canada, and, in the Province of Quebec, the Collège des médecins du Québec. The accreditation process cycles throughout a 6-year period, culminating with a regular, on-site, external review (ER) performed by a team of representatives from the 3 colleges.<sup>1-6</sup> The colleges mandate that, 2 years before the ER, some form of internal review (IR) be performed to ensure that training standards are being met and that

programs will be prepared for the on-site ER. The colleges do not mandate the specifics of the IR process.<sup>2,3,6</sup>

Currently, there is discussion at the national level about restructuring the accreditation process and possibly exempting certain programs from the regular, on-site ER. In the United States, there has already been a restructuring of the accreditation process by the Accreditation Council for Graduate Medical Education (ACGME) with the New Accreditation System (NAS),<sup>7</sup> which will likewise result in fewer face-to-face meetings. We believe that the IR must be rigorous and that the postgraduate medical education (PGME) office must demonstrate to the accreditation authorities that a robust IR was undertaken.

Currently, there is no standardized process to IRs in Canada. The IR varies across the 17 Faculties of Medicine, and to our knowledge, none simulate all aspects of an ER for their IR. There is neither published research (databases searched included Ovid, Embase, and PubMed) describing possible processes of conducting an IR, nor are there any published reports on the perceptions of key participants of the process and the outcome. One author suggested that the IR should be rigorous and tough.<sup>8</sup> We feel that a standardized, rigorous, and comprehensive IR process would appropriately identify program strengths and weaknesses. The purpose of our study was to simulate all

All authors are with McGill University, Montreal, Quebec, Canada. **Kenneth Doyle, MD**, is Director of Accreditation, Postgraduate Medical Education Office; **Meredith Young, PhD**, is Assistant Professor, Centre for Medical Education; and **Sarkis Meterissian, MDCM**, is Associate Dean, Postgraduate Medical Education Office and the Centre for Medical Education.

Funding: The authors report no external funding source for this study.

Corresponding author: Kenneth Doyle, MD, Postgraduate Medical Education, Faculty of Medicine, McGill University, 3655 Promenade Sir William Osler, Montreal, Quebec, H3G 1Y6, Canada, fax 514.398.3595, kenneth.doyle@mcgill.ca

Received January 23, 2013; revisions received May 27, 2013, and July 20, 2013; accepted August 14, 2013.

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00026.1>

aspects of a regular, on-site ER as part of our IR process and to assess the perceptions of both the process and the outcome of the IR process.

## Methods

This study was undertaken at McGill University in Montreal, Canada. Forty-two Royal College programs and 7 family medicine programs were reviewed during a 4-day period.

## Participants

Participants included 53 program directors (PDs), 45 program coordinators (PCs), 30 faculty surveyors, and 15 resident surveyors (RSs). Surveyors were recruited from within the faculty and included individuals with previous national ER experience (12), former PDs, or those with leadership experience and/or a demonstrated interest in medical education. The RSs were recruited with the assistance of the Association of Residents of McGill and included those with national ER experience (1), those with expressed interest in education, or those with chief resident experience. Seven PDs volunteered to participate as secondary surveyors (to assist the faculty surveyor with the review of a program other than their own). All surveyors were assigned to programs eliminating any possible conflict of interest (ie, surveying programs in their own discipline or those of close colleagues).

## Preparation and Training

All surveyors, RSs, PDs, and PCs underwent training identical to the workshops provided by the 3 colleges to new PDs, surveyors, and PCs.

The PGME office distributed program-specific documents to each of the programs, containing the presurvey questionnaire, the general standards of accreditation,<sup>2-4,6</sup> the most recent ER accreditation report, survey templates, and any relevant, program-specific documents.<sup>9</sup> With the exception of the Specialty Committee Report (a report including specialty-specific issues that the survey teams must attend to, which is generated by a national committee that reviews their specialty's respective presurvey questionnaire), the materials sent to the programs were identical to those that would have been sent during a regular, on-site ER. Our simulation pertains to the *regular* ER that occurs at the end of the 6-year accreditation cycle, not the *mandated* ER, which occurs when specific concerns are identified following a regular, on-site ER.

## Survey Week

All aspects of the survey week schedule and the individual program survey schedules were identical to an on-site ER, including all the required meetings (document review,

### What was known

Graduate medical education accreditation in Canada, as in other nations, seeks to reduce burden while maximizing the efficacy and value of reviews.

### What is new

Simulation of all aspects of a regular, on-site, external review using reviewers internal to the institution to create a robust, outcomes-focused, internal review process.

### Limitations

Single institution, lack of a control group, and nonvalidated instrument for assessing acceptability and utility of the simulated review process may limit generalizability.

### Bottom line

A rigorous internal review could replace external reviews for stable, high-performing programs, allowing external reviews to focus on programs with less than full approval following an internal review.

meetings with the PD, department chair, faculty, residents, and the Residency Program Committee) as well as the sequence and duration of these meetings. The survey team convened in the evening to discuss the programs reviewed each day and to recommend a category of accreditation. These meetings were chaired by 1 of the authors (K.D.). A 15-minute exit meeting between the surveyor and the PD was held for each program on the following morning.

## Post-IR Follow-Up

The post-IR follow-up period included the following: (1) an action plan submitted by each program to the PGME office indicating how the cited weaknesses would be addressed; (2) a meeting between the associate dean of PGME and the PDs of programs that did not receive full approval to discuss how the cited weaknesses would be addressed; (3) a meeting was held, for programs receiving "provisional approval with internal review," between the associate dean of PGME and the Residency Program Committees 1 year after the IR to review progress with the action plan; and (4) a repeat review was held for programs receiving either "provisional approval with external review" or "notice of intent to withdraw accreditation," 1 year after the IR, with invited, experienced, specialty-specific surveyors from outside the McGill network as well as RSs from within McGill.

## The Survey

To assess perceptions of the process and outcome, 4 parallel surveys were developed (1 each for PDs, surveyors [both primary and secondary], RSs, and PCs). The McGill Postgraduate Medicine Education Committee reviewed the surveys for content, but no other testing was done.

The survey was distributed using an online survey tool and sent to 53 PDs, 30 primary surveyors, 7 secondary

TABLE 1 PERCEPTION OF HOW PREPARED EACH GROUP FELT BEFORE THE INTERNAL REVIEW

|                                                                     | Program Director,<br>n = 36 | Surveyor,<br>n = 30 | Resident Surveyor,<br>n = 12 | Program<br>Coordinator, n = 32 |
|---------------------------------------------------------------------|-----------------------------|---------------------|------------------------------|--------------------------------|
| Mean (SD)                                                           | 4.6 (0.87)                  | 5.0 (0.72)          | 4.9 (0.79)                   | 4.9 (1.17)                     |
| Very well or extremely well prepared (5 or 6), <sup>a</sup> No. (%) | 22 (61)                     | 25 (83)             | 10 (83)                      | 22 (69)                        |
| Satisfactory or well prepared (3 or 4), <sup>a</sup> No. (%)        | 14 (39)                     | 5 (17)              | 2 (17)                       | 9 (28)                         |
| Not well prepared (1 or 2), <sup>a</sup> No. (%)                    | 0                           | 0                   | 0                            | 1 (3)                          |

<sup>a</sup> Likert-type scale: 1, not well prepared at all; 2, not well prepared; 3, satisfactory; 4, well prepared; 5, very well prepared; 6, extremely well prepared.

surveyors, 15 RSs, and 45 PCs, covering 49 programs (there are more PDs than programs because some programs have co-PDs). The initial invitation to complete the survey was sent 2 weeks postreview, and the last reminder occurred 4 months later. Up to 4 reminders were sent. The survey contained both open-ended items and 6-point Likert-style rating items (with 1 being the lowest and 6 being the highest or best possible response, anchored as appropriate). Free response sections were included following each question.

The project received ethics approval from the McGill Faculty of Medicine Institutional Review Board.

### Data Analysis

Survey responses were analyzed descriptively using means and weighted means, and 1-way ANOVA and *t* tests were used for comparison across respondent groups where appropriate. Free response comments were analyzed for prevailing themes.

## Results

### Participants

The response rate was 73%, ranging from 81% (30 of 37) of surveyors, 80% (12 of 15) of RSs, 71% (32 of 45) of PCs, and 67% (35 of 52) of PDs.

### Surveyor Experience

Of the 30 responding surveyors, 12 had previous survey experience at the college level, with 9 having participated in 3 or more surveys outside our university. Nine of the surveyors had experience with previous IRs. Two free-text comments explicitly compared our IR to their experience reviewing at the college level. One commented our IR was “better organized” with “more pretraining,” and the other noted that the “reviewers were more efficient than the usual Royal College reviewers.”

### Perceptions of Preparation

Most of the surveyors and RSs felt they were either very well or extremely well prepared for the survey process, whereas fewer PDs felt similarly (TABLE 1). There were no significant differences between the 4 groups in their perceptions of how well prepared they were for the internal review ( $F_{3,106} = 1.20$ ,  $P = .31$ ). One PC felt poorly prepared for the IR, having just returned from a leave of absence.

### Perceptions of the Simulation

Among the surveyors, RSs, PDs, and PCs who had previously participated in ERs, most perceived that the IR simulated an actual review either extremely or very well (TABLE 2). There were no significant differences between the 4 groups of participants ( $F_{3,50} = 1.27$ ,  $P = .30$ ).

TABLE 2 HOW WELL OUR INTERNAL REVIEW SIMULATED AN EXTERNAL REVIEW

|                                                                    | Program Director, n = 14 | Surveyor, n = 17 | Resident Surveyor, n = 5 | Program Coordinator, n = 21 |
|--------------------------------------------------------------------|--------------------------|------------------|--------------------------|-----------------------------|
| Mean, of 6, (SD)                                                   | 4.6 (1.12)               | 4.76 (0.90)      | 5.0 (0)                  | 5.14 (0.72)                 |
| Extremely or very well simulated (5 or 6), <sup>a</sup> No. (%)    | 8 (57)                   | 12 (71)          | 5 (100)                  | 19 (90)                     |
| Satisfactory (3 or 4), <sup>a</sup> No. (%)                        | 5 (36)                   | 5 (29)           | 0                        | 2 (10)                      |
| Poorly or not at all well simulated (1 or 2), <sup>a</sup> No. (%) | 1 (7)                    | 0                | 0                        | 0                           |

<sup>a</sup> Likert-type scale: 1, not at all well simulated; 2, poorly simulated; 3, satisfactorily simulated; 4, well simulated; 5, very well simulated; 6, extremely well simulated.

TABLE 3 PERCEPTIONS OF RIGOR

|                                                               | Program Directors,<br>n = 35 | Surveyors, n = 30 | Resident Surveyors,<br>n = 12 |
|---------------------------------------------------------------|------------------------------|-------------------|-------------------------------|
| Mean, of 6, (SD)                                              | 4.97 (1.04)                  | 5.30 (0.60)       | 5.33 (0.51)                   |
| Extremely or very rigorous (5 or 6), <sup>a</sup> No. (%)     | 26 (74)                      | 28 (93)           | 11 (92)                       |
| Satisfactorily rigorous (3 or 4), <sup>a</sup> No. (%)        | 9 (26)                       | 2 (7)             | 1 (8)                         |
| Minimal or not rigorous at all (1 or 2), <sup>a</sup> No. (%) | 0                            | 0                 | 0                             |

<sup>a</sup> Likert-type scale: 1, not at all rigorous; 2, minimally rigorous; 3, satisfactorily rigorous; 4, rigorous; 5, very rigorous; 6, extremely rigorous.

Most of the PDs, surveyors, and RSs perceived the IR process to be either extremely or very rigorous (TABLE 3). There was no difference found between perceptions of rigor across the 3 study populations ( $F_{2,78} = 2.31$ ,  $P = .10$ ). One PD (from a program that did not receive full approval) reported that the IR was “overly harsh.”

Most of the PDs, surveyors, and RSs perceived the IR process to be either extremely or very unbiased (TABLE 4). No significant interaction was found between ratings of bias and participant group ( $F_{2,75} = 2.06$ ,  $P = .13$ ). Planned comparisons found a trend between PDs and RSs, in which residents found the process to be less unbiased than PDs found it ( $t_{45} = 1.98$ ,  $P = .05$ ).

Three of the 4 free-response comments expressed concern about programs being reviewed by close colleagues (the programs supervised by those PDs all received approval). One respondent noted that the surveyor was “purposefully tough,” while commenting that it is difficult in a small university environment to avoid this bias. However, respondents perceived the IR process to be very unbiased (5 of 6). The fourth respondent perceived the IR process as biased (2 of 6), in favor of listening to residents and that the residents were guided to speak negatively about the program. No prevailing themes were identified in the comment section. One surveyor commented that we are often more critical of ourselves, but that could allow important changes to take place.

Most of the PDs, surveyors, and RSs perceived the IR process to be either extremely or very fair (TABLE 5). A significant interaction was found between ratings of fairness and the 3 groups ( $F_{2,74} = 4.31$ ,  $P = .02$ ). Post hoc analyses found a difference between PDs and RSs, where the PDs found the process to be less fair than RSs did (mean RS = 5.58; mean PD = 4.76;  $t_{44} = 2.59$ ;  $P < .01$ ). No prevailing themes were identified in the comment section.

#### Assessment of Program Strengths and Weaknesses

Most of the PDs perceived that the cited program strengths reflected extremely or very well their program strengths (TABLE 6). Only 1 PD (from a program receiving “provisional approval with external review”) thought that the cited strengths did not reflect program strengths. Most of the PDs perceived that the cited list of weaknesses reflected their program weaknesses well.

In the free response section, no prevailing themes were identified. Three PDs commented on cited program strengths. Two PDs (both from programs that received a recommended accreditation status of “provisional approval with internal review”) commented that important program strengths may have been overlooked, perhaps because too much emphasis was placed on resident complaints while disregarding the academic input and effort of the staff or overlooking program specifics because of the lack of specialty committee input. For the cited weaknesses, 8 PDs (4 from programs receiving approval) commented that

TABLE 4 PERCEPTIONS OF BIAS

|                                                               | Program Directors,<br>n = 35 | Surveyors, n = 30 | Resident Surveyors,<br>n = 12 |
|---------------------------------------------------------------|------------------------------|-------------------|-------------------------------|
| Mean (SD)                                                     | 4.63 (1.19)                  | 4.83 (0.99)       | 5.33 (0.47)                   |
| Extremely or very unbiased (5 or 6), <sup>a</sup> No. (%)     | 23 (66)                      | 23 (77)           | 12 (100)                      |
| Satisfactorily unbiased (3 or 4), <sup>a</sup> No. (%)        | 9 (26)                       | 5 (17)            | 0                             |
| Minimal or not at all unbiased (1 or 2), <sup>a</sup> No. (%) | 3 (9)                        | 2 (7)             | 0                             |

<sup>a</sup> Likert-type scale: 1, not at all unbiased; 2, minimally unbiased; 3, satisfactorily unbiased; 4, unbiased; 5, very unbiased; 6, extremely unbiased.

TABLE 5 PERCEPTIONS OF FAIRNESS

|                                                           | Program Directors,<br>n = 34 | Surveyors, n = 26 | Resident Surveyors,<br>n = 12 |
|-----------------------------------------------------------|------------------------------|-------------------|-------------------------------|
| Mean, of 6, (SD)                                          | 4.76 (1.04)                  | 5.13 (0.82)       | 5.58 (0.51)                   |
| Extremely or very fair (5 or 6), <sup>a</sup> No. (%)     | 24 (71)                      | 22 (85)           | 12 (100)                      |
| Satisfactorily fair (3 or 4), <sup>a</sup> No. (%)        | 9 (26)                       | 4 (15)            | 0                             |
| Minimal or not at all fair (1 or 2), <sup>a</sup> No. (%) | 1 (3)                        | 0                 | 0                             |

<sup>a</sup> Likert-type scale: 1, not at all fair; 2, minimally fair; 3, satisfactorily fair; 4, fair; 5, very fair; 6, extremely fair.

weaknesses and areas for improvement were appropriately identified, that some surveyors were overly harsh and some not harsh enough, and that some surveyors focused too much on resident complaints.

### Survey Findings in Relation to Review Outcome

When the results were stratified based on review outcome, the perception of PDs varied (TABLE 7). The PDs whose programs did not receive full approval considered the IR process to be less fair ( $t_{32} = 2.93$ ,  $P = .006$ ), more biased ( $t_{32} = 3.24$ ,  $P < .005$ ), and did not accurately reflect program strengths ( $t_{32} = 3.52$ ,  $P < .005$ ) and weaknesses ( $t_{32} = 2.66$ ,  $P = .05$ ) compared with PDs whose programs received a positive review. However, those PDs perceived they were well prepared for the IR, ( $t_{32} = 1.67$ ,  $P = .11$ ), and considered the IR process to be rigorous ( $t_{32} = 0.399$ ,  $P = .69$ ).

### Discussion

We successfully simulated a tri-college, on-site ER for our IR process. Our IR contained several key elements that are not routinely included, such as (1) training of all participants using the same workshops as provided by the 3 colleges, (2) program review conducted by physicians from another specialty, (3) team discussion about program strengths and weaknesses, (4) recommendations of an actual category of accreditation, (5) a resident report and

resident participants as surveyors, and (6) structured, post-IR, follow-up procedures. We feel that these elements added a degree of rigor to our IR process.

Overall, the participants (surveyors, PDs, RSs, and PCs) felt well prepared for the IR. The process was perceived as rigorous, fair, unbiased, and the cited program strengths and weaknesses were appropriate. The PDs whose programs received less favorable program reviews viewed our IR process less favorably than others did.

Our study has several limitations. One is the lack of a control group because this design may limit our ability to quantify the direct effect of a simulated ER on actual ER outcomes. However, we believe that the predicted pedagogic value of a rigorous IR did not permit us to exclude certain programs (and therefore create a control group) from this review process. Another limitation is the lack of previously validated tools to assess the views of the participants about the acceptability of the IR. Because of the lack of previously published tools, we relied on a locally developed and reviewed survey tool. It is likely that there was variability in how participants interpreted and interacted with the survey items. A third limitation is the lack of a current national specialty committee report for the IR. During a regular, on-site ER, members of each specialty committee review the presurvey questionnaire and subsequently derive a list of concerns that need to be clarified during the visit. Although this is not a limitation based on

TABLE 6 PROGRAM DIRECTORS' PERCEPTION OF HOW WELL THE REVIEW PROCESS REFLECTED OVERALL PROGRAM STRENGTHS AND WEAKNESSES (N = 35)

|                                                                                             | Strengths   | Weaknesses  |
|---------------------------------------------------------------------------------------------|-------------|-------------|
| Mean (SD)                                                                                   | 4.77 (1.00) | 4.51 (0.85) |
| Reflects strengths or weaknesses extremely well or very well (5 or 6), <sup>a</sup> No. (%) | 26 (74)     | 18 (51)     |
| Reflects strengths or weaknesses satisfactorily (3 or 4), <sup>a</sup> No. (%)              | 8 (23)      | 17 (49)     |
| Reflects strengths or weaknesses poorly or not at all (1 or 2), <sup>a</sup> No. (%)        | 1 (3)       | 0           |

<sup>a</sup> Likert-type scale: 1, does not reflect program strengths and weaknesses at all; 2, poorly reflects strengths and weaknesses; 3, satisfactorily reflects strengths and weaknesses; 4, reflects strengths and weaknesses; 5, reflects strengths and weaknesses very well; 6, reflects strengths and weaknesses extremely well.

| TABLE 7 PROGRAM DIRECTOR PERCEPTIONS BASED ON REVIEW OUTCOME                  |             |             |             |             |             |             |
|-------------------------------------------------------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                                                                               | Strengths   | Weaknesses  | Rigor       | Unbiased    | Fair        | Prepared    |
| Approval, Mean (SD)                                                           | 5.13 (0.63) | 4.91 (0.84) | 4.96 (0.93) | 5.00 (0.78) | 5.00 (0.90) | 4.74 (0.62) |
| Provisional approval or notice of intent to withdraw accreditation, Mean (SD) | 4.00 (1.26) | 4.09 (0.83) | 4.82 (0.98) | 3.81 (1.32) | 3.91 (1.22) | 4.27 (0.96) |
| P value                                                                       | .001        | .01         | .69         | .003        | .006        | .11         |

study design, the responses generated to the concerns raised by the specialty committee could potentially have an influence on the assigned accreditation status. It was not possible to include that component in our initial attempt to simulate the ER process.

An interesting, but not unexpected, finding was the dependence of PDs’ perception of the process on the recommended accreditation status. This is important for all accrediting bodies intending to assess their review process and fits well with the literature regarding receiving and internalizing negative feedback.<sup>10</sup>

Following a regular, on-site ER, the colleges require programs not receiving full approval to undergo either a mandated IR or an ER within 2 years, depending on the review outcome. However, none of the colleges specify the exact nature of the follow-up required after an IR.<sup>3,6</sup> We required all our programs to participate in follow-up, regardless of accreditation status, by submitting an action plan to the PGME office outlining how the cited weaknesses would be addressed. Further, programs that received less than full approval had either a follow-up IR or an ER within 1 year of the IR. We believe that this follow-up process enhanced the simulation of a regular, on-site ER and added a further degree of structure and rigor to our IR process.

It was not our intent to study the costs of running a rigorous IR. The costs paid by our PGME office for the IR were essentially the same as those for an ER because our total budget was the same for both.

In the United States, the ACGME has already implemented the NAS in 7 core specialties. Under the NAS, the midcycle internal reviews will disappear (replaced by an annual review of accreditation and stepped-up expectations for the required internal annual program evaluation), and on-site reviews will be replaced by a comprehensive self-study (scheduled at 10-year intervals) unless the data review suggests a problem; in which case, a site visit is triggered. Consequently, the annual internal program evaluation required by the ACGME will need to be structured and rigorous.

Conclusion

We were able to simulate a regular, on-site, external review, and most of the respondents perceived the process to be rigorous, fair and unbiased, and the cited lists of program strengths and weaknesses were appropriate. We postulate that future on-site ERs could be limited to programs that receive less than full approval following a structured and rigorous IR. Further analysis comparing the outcomes of an IR and an on-site ER done by the external accreditor is needed.

References

1 Royal College of Physicians and Surgeons of Canada. Residency accreditation promotes excellence. <http://www.royalcollege.ca/portal/page/portal/rc/credentials/accreditation>. Published 2011. Accessed August 1, 2012.

2 Royal College of Physicians and Surgeons of Canada. General standards of accreditation applicable to all residency programs: B standards. [http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/accreditation\\_blue\\_book\\_b\\_standards\\_e.pdf](http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/accreditation_blue_book_b_standards_e.pdf). Published July 2011. Accessed August 1, 2012.

3 Royal College of Physicians and Surgeons of Canada. General standards applicable to the university and affiliated sites: A standards. [http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/accreditation\\_purple\\_book\\_a\\_standards\\_e.pdf](http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/accreditation_purple_book_a_standards_e.pdf). Published July 2011. Accessed September 15, 2009.

4 Royal College of Physicians and Surgeons of Canada. General information concerning accreditation of residency programs. [http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/genaccred\\_e.pdf](http://www.royalcollege.ca/portal/page/portal/rc/common/documents/accreditation/genaccred_e.pdf). Published June 2006. Accessed September 15, 2009.

5 College of Family Physicians of Canada. Residency program accreditation. [http://www.cfpc.ca/Residency\\_Program\\_Accreditation/](http://www.cfpc.ca/Residency_Program_Accreditation/). Published 2006. Accessed September 15, 2009.

6 College of Family Physicians of Canada. General information concerning accreditation of residency programs. <http://www.cfpc.ca/uploadedFiles/Red%20Book%20English.pdf>. Accessed September 15, 2009.

7 Nasca Thomas J, Philbert I, Brigham T, Flynn T. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1056.

8 Walsh A. A survivor’s guide to the accreditation cycle. <http://www.cfpc.ca/uploadedFiles/Education/Survivor%20Guide.pdf>. Published 2006. Accessed September 15, 2009.

9 Royal College of Physicians and Surgeons of Canada. Information by discipline: information by specialty and information by subspecialty. [http://www.royalcollege.ca/rc/faces/oracle/webcenter/portalapp/pages/ibd.jsp;jsessionid=AETZB1zuVANSnYBngWd\\_YHDb077oqMCfdKsGbk6TsAffPVXkyFI-1538745610?lang=en&\\_afLoop=4381892891374674&\\_afWindowMode=o&\\_afWindowId=null#%40%3F\\_afWindowId%3Dnull%26\\_afLoop%3D4381892891374674%26lang%3Den%26\\_afWindowMode%3D%26\\_adf.ctrl-state%3Dn7uifuh91\\_4](http://www.royalcollege.ca/rc/faces/oracle/webcenter/portalapp/pages/ibd.jsp;jsessionid=AETZB1zuVANSnYBngWd_YHDb077oqMCfdKsGbk6TsAffPVXkyFI-1538745610?lang=en&_afLoop=4381892891374674&_afWindowMode=o&_afWindowId=null#%40%3F_afWindowId%3Dnull%26_afLoop%3D4381892891374674%26lang%3Den%26_afWindowMode%3D%26_adf.ctrl-state%3Dn7uifuh91_4). Published 2011. Accessed August 1, 2012.

10 Sargeant J, Mann K, Ferrier S. Exploring family physicians’ reactions to multisource feedback: perceptions of credibility and usefulness. *Med Educ*. 2005;39(5):497–504.