



**DALHOUSIE**  
*University*

Continuing Medical Education  
Clinical Research Centre  
5849 University Avenue  
Halifax, Nova Scotia  
Canada B3H 4H7

**FINAL REPORT: A PILOT STUDY OF THE ALBERTA PHYSICIAN ACHIEVEMENT  
REVIEW (PAR) PROGRAM IN NOVA SCOTIA**

Presented by:  
NS PAR Research Working Group  
Office of Continuing Medical Education  
Faculty of Medicine  
Dalhousie University

Presented to:  
The College of Physicians and Surgeons of Nova Scotia

April 21, 2003

**NS PAR Research Working Group members:**

*Continuing Medical Education*

Joan Sargeant MEd, Director, Program Development & Evaluation (Chair)

Suzanne Ferrier MSc, Research Associate

Vonda Hayes MD, Director, Physician Enhancement Program

Douglas Sinclair MD, Associate Dean

*Community Health and Epidemiology*

Donald Langille MD, Associate Professor

*Division of Medical Education*

Karen Mann PhD, Director

*Family Physician Representative*

Philip Muirhead MD, Musquodoboit Harbour

## TABLE OF CONTENTS

<b>List of tables</b> .....	<b>3</b>
<b>List of appendices</b> .....	<b>5</b>
<b>Executive summary</b> .....	<b>6</b>
<b>Introduction</b> .....	<b>9</b>
<b>Methods</b> .....	<b>9</b>
Recruitment.....	9
Data collection.....	10
Analysis.....	14
<b>Results</b> .....	<b>16</b>
Response rates .....	16
Research question #1.....	19
Research question # 2.....	23
Limitations of the study.....	33
<b>Conclusions and discussion</b> .....	<b>34</b>
Research question #1:.....	34
Research question #2:.....	34
<b>Recommendations</b> .....	<b>38</b>
<b>References</b> .....	<b>39</b>

## LIST OF TABLES

<b>Table 1:</b> Data collection methods to address the two NSPAR research questions.....	10
<b>Table 2:</b> Response rates for NSPAR pilot project.....	16
<b>Table 3:</b> Demographic information of three physician populations: NSPAR volunteers who participated in the pilot study, NSPAR volunteers who completed the Dalhousie evaluation questionnaire, and NS physicians.....	17
<b>Table 4:</b> Physicians' reasons for participating in NSPAR as reported on the physician evaluation questionnaire .....	17
<b>Table 5:</b> Demographic information of three patient study populations: Patients responding to the PAR questionnaire (N=3442), the Dalhousie CME questionnaire (N =2223), and participating in the interview (N=591).....	18
<b>Table 6:</b> Number (%) of participants providing requested number of names of reviewers, receiving required number of responses, and mean, range of respondents per participant, by reviewer group.....	19
<b>Table 7:</b> Descriptive statistics and reliability coefficients of the 4 PAR questionnaires for 142 Nova Scotia physicians.....	20
<b>Table 8:</b> Descriptive data for the PAR attributes measured by the SAQ.....	20
<b>Table 9:</b> Descriptive data for the PAR attributes measured by medical colleagues.....	21
<b>Table 10:</b> Descriptive data for the PAR attributes measured by co-workers.....	21
<b>Table 11:</b> Descriptive data for the PAR attributes measured by patients.....	22
<b>Table 12:</b> Reliability coefficients for PAR questionnaires and attributes within the questionnaires .....	23
<b>Table 13:</b> Number (%) of “unable to assess” (UA) responses by over 10% of medical colleagues, by item, type of colleague and attribute.....	25
<b>Table 14:</b> Number (%) of “unable to assess” (UA) responses selected by over 10% of co-workers and patients, by item and attribute.....	27
<b>Table 15:</b> Reviewers' responses to whether or not they knew the physician well enough to answer most of the questions adequately.....	27
<b>Table 16:</b> Medical colleagues' and coworkers' familiarity with physician participant.....	28
<b>Table 17:</b> Medical colleagues' and co-workers' familiarity ratings and the mean questionnaire score given to the NSPAR participant.....	28
<b>Table 18:</b> Reviewers' self-reported levels of comfort in participating in NSPAR.....	29

**Table 19:** Reviewers' self-reported ability to answer the NSPAR questions truthfully.....29

**Table 20:** Percentage of physician participants agreeing with the following descriptors about the NSPAR recruitment process and results, for medical colleagues, coworkers, and patients..... 30

**Table 21:** Relationship between physicians' overall mean MCAQ scores from their medical colleagues and their agreement ratings on three of the Dal evaluation questionnaires .....31

**LIST OF APPENDICES**

**Appendix A:** Physician evaluation questionnaire.....40

**Appendix B:** Medical colleagues and co-worker evaluation questionnaire..... 44

**Appendix C:** Patient evaluation questionnaire..... 46

**Appendix D:** Patient interview questions..... 48

**Appendix E:** Focus group questions.....51

**Appendix F:** Descriptive data for each SAQ item organized by attribute..... 53

**Appendix G:** Descriptive data for each MCAQ item organized by attribute.....55

**Appendix H:** Descriptive data for each CWAQ item organized by attribute..... 57

**Appendix I:** Descriptive data for each PAQ item organized by attribute..... 58

**Appendix J:** Sample physician NSPAR report..... 60

## **EXECUTIVE SUMMARY**

The College of Physicians and Surgeons of Nova Scotia (CPSNS) wished to undertake a pilot study of the College of Physicians and Surgeons of Alberta (CPSA) Physician Achievement Review (PAR) program. PAR assesses physician performance using a multi-source feedback approach which includes responses from physicians themselves, their medical colleagues, patients, and co-workers. The goal of PAR is to enhance physician performance using the principles of quality improvement. CPSNS invited Dalhousie University Office of Continuing Medical Education to conduct the pilot study of PAR for Nova Scotia family physicians (NSPAR).

The specific research questions addressed by the study were:

1. How do Nova Scotia physicians, their peers, co-workers and patients respond to a set of instruments developed in Alberta designed to assess physician performance?
2. Is it feasible to use the PAR instruments and processes in Nova Scotia?

Data collection methods for Research Question #1 included the administration of the PAR questionnaires for self, medical colleagues, coworkers and patients, and compilation of individual physician participant reports based upon these results. For Research Question #2, methods included completion of post-survey evaluation questionnaires by physician participants, medical colleagues, coworkers and patients, and patient interviews and physician participant focus groups.

Although a sample population of 100 family physicians was needed for the study, over 170 volunteered. Of these, 167 initiated the study and 142 (87.7%) completed the process and received NSPAR reports.

### **Summary of key conclusions**

1. The vast majority (88.5%) of physician participant respondents to the written post-survey evaluation felt that the NSPAR report provided useful information, and 61% indicated that they had made or were planning to make practice changes in response to it. This is generally positive feedback on the usefulness of the NSPAR process. Of the 75 intended changes identified, almost 75% addressed communication issues, most with patients but some also with medical colleagues and coworkers.
2. Overall, most physician participants agreed that NSPAR was useful as a process for providing NS physicians with a general overview of their practice. Strengths of NSPAR are: the inclusion of patients, the broad perspective which it provides on one's practice, and the inclusion of Nova Scotia physicians in a regular and ongoing evaluation and feedback process. Areas for improvement include lack of specificity of results which diminishes its usefulness as an educational tool, and concerns that medical colleague feedback is of less value than that of other reviewers.
3. While the item means on the medical colleague, coworker and patient questionnaires were generally high (above 4.0 on a 5-point likert scale, for all but 3 of the total 88 items), analysis of the results showed that medical colleagues and coworkers who knew the physician participants well or very well gave significantly higher scores than those who knew them somewhat well or not well.

4. Medical colleagues more frequently indicated they were “unable to assess” the physician than other reviewer groups, and consultant physicians responded “unable to assess” much more frequently than peers.
5. A large proportion of medical colleagues (44%) reported having difficulty answering specific questions, as compared to coworkers (24%) and patients (7%). The most common reason given for this was that the question was, “not applicable to my relationship with the physician”.
6. While 91% of physician participants agreed that medical colleagues should be included as reviewers and 79% felt comfortable recruiting them, only 65.5% agreed with their results. Focus group respondents provided some explanations for this.
7. Focus group participants also reported that receiving negative feedback; i.e., being rated below the mean or receiving an “information flag”, was troubling for them. This suggests a need for delivering the feedback in a sensitive and careful manner, as well as attempting to ensure the participation of reviewers who know physicians sufficiently well.
8. If NSPAR were to be accepted, physicians request clear and open communication by the College about the purpose of NSPAR and use of NSPAR data.

### **Recommendations**

If the CPSNS were to adopt the PAR program, we would make the following recommendations:

#### *Communication and support*

1. Communication with membership is fundamental to the success of new initiatives, as demonstrated by the College’s successful NSPAR pre-pilot communication process. If NSPAR were accepted, we would again encourage clear, open and consistent communication with members about its rationale, purpose, processes and confidentiality procedures.
2. Although many respondents reported finding the report useful, some had difficulty interpreting the results and understanding how they might improve their practice. We would recommend the provision of a confidential counseling service to assist physicians with understanding and interpreting their reports, identifying ways to improve their practice, and accessing resources for continuing professional development.

#### *Medical colleague reviewers*

3. As the medical colleague reviewer group appeared to respond differently than the other groups; i.e., appeared to be less familiar with the physicians, had more difficulty responding to items, and were unable to respond to more items than other groups, we suggest the following:
  - 3.1 Conduct further analyses of medical colleague responses and results to determine if other factors influence both their ability to respond and physician participants’ responses to their results; e.g., geographic or practice type factors. This is important as few studies have been conducted in this area.

- 3.2 Consider including a means of adjusting results to reflect level of familiarity with the physician.
- 3.3 Consider excluding items which a larger percentage of medical colleague reviewers are “unable to assess”, or assessing them in a different way; e.g., those relating to psycho-social care. The latter approach may be preferable. However, this would change the assessment tool and would require further study.
- 3.4 As consultant medical colleagues reported more “unable to assess” items than peer colleagues, consider a means for identifying consultants who know physicians sufficiently well to complete the items on the questionnaire. It may also be helpful to review the separate consultant questionnaire used in the original Alberta PAR pilot study (Hall, 1999), and discuss its use with the Alberta research team.

Finally, regardless of the College’s decision to adopt NSPAR, we recommend providing a summarized report of pilot study findings to NSPAR physician participants. They expressed much interest throughout the process and appear sincerely interested in its outcome.

## **INTRODUCTION**

The College of Physicians and Surgeons of Nova Scotia (CPSNS) wished to undertake a pilot study of the College of Physicians and Surgeons of Alberta (CPSA) Physician Achievement Review (PAR) program. This program assesses aspects of physician performance using a multi-source feedback approach, including responses from physicians themselves, their medical colleagues, patients, and co-workers. The goal of the PAR program is to enhance physician performance using a quality improvement process. Results of a pilot study of the Alberta PAR program indicated high statistical validity and reliability of the instruments. Results also demonstrated that physicians participating in the pilot considered or had implemented changes in their practices on the basis of the feedback they had received from the PAR program. (Violato, 1997; Hall, 1999; Fidler, 1999). On the basis of this work, CPSNS invited Dalhousie University Office of Continuing Medical Education (Dalhousie CME) to undertake a pilot study of Nova Scotia family physicians, to test the instruments and processes developed in the PAR program and evaluate its feasibility for Nova Scotia (NSPAR).

The specific research questions addressed by this study were:

1. How do Nova Scotia physicians, their peers, co-workers and patients respond to a set of instruments developed in Alberta designed to assess physician performance?
2. Is it feasible to use the PAR instruments and processes in Nova Scotia?

## **METHODS**

### **Recruitment**

The goal was to recruit 100 family practitioners who had been in active office-based general practice in Nova Scotia for at least five years. Recruitment for the pilot study was preceded by the CPSNS' NSPAR Communication Strategy, which involved a focus group and province-wide presentations, described below. Dalhousie CME conducted the recruitment through invitation letters mailed to eligible physicians. As an incentive for their participation, physician volunteers received a 5-year exemption from the Atlantic Provinces Medical Peer Review (APMPR). Additionally, volunteers were eligible to receive CME credit for their participation through the College of Family Physicians of Canada (MAINPRO-C or MAINPRO-M1)

### *CPSNS Communications Strategy*

In September, 2001 the CPSNS held a PAR focus group with influential family practitioners, including current and former executive members of the CPSNS, the Medical Society of Nova Scotia and the Nova Scotia College of Family Physicians. The purpose of this focus group was to seek input from physicians for the communications package and their endorsement of it. Revisions were made to the presentation as a result of the feedback provided during the focus group.

The CPSNS mailed an NSPAR information package to all members and created a NSPAR website, which provided information about the pilot study, posted frequently-asked questions and samples of the PAR questionnaires. In October and November, 2001, CPSNS conducted 19 information sessions around the province. These included a presentation about the pilot study, conducted by a physician opinion leader known within the area, followed by a question and answer session. The CPSNS Registrar and/or Deputy Registrar were available to answer

questions. At each session, volunteer forms for the NSPAR pilot study were available for interested physicians to complete and fax to Dalhousie CME.

### Invitation letters

An invitation letter and consent form were mailed from Dalhousie CME to all family practitioners in NS who met the eligibility criteria (i.e., in active practice for a minimum of five years in NS). This letter explained the purpose and procedure of the NSPAR pilot project. Interested physicians were requested to complete the consent form and return it to Dalhousie CME to indicate their willingness to participate in the study. A reminder letter was sent to physicians who had not responded to the invitation, approximately 3 weeks following the initial mailout.

### Data collection

Two research questions were addressed by the NSPAR project:

1. How do Nova Scotia physicians, their peers, co-workers and patients respond to a set of instruments developed in Alberta designed to assess physician performance?
2. Is it feasible to use the PAR instruments and processes in Nova Scotia?

Table 1 shows the data collection methods and responsibility for each question. Customer Information Systems Inc. (CIS, an independent research agency specializing in evaluation services in Edmonton, Alberta, was contracted by the College of Physicians and Surgeons of Alberta (CPSA) to administer the Alberta PAR program. CPSNS contracted with CIS to administer the PAR tools for the pilot project in Nova Scotia, to aid in answering the first research question. The research team at Dalhousie CME managed the data collection for the second research question.

**Table 1:** Data collection methods to address the two NSPAR research questions.

Research Question	Data collection tools for each reviewer group				Responsible for data collection
	Family/General Physician	Medical Colleague	Co-worker	Patient	
1	PAR questionnaire	PAR questionnaire	PAR questionnaire	PAR questionnaire	CIS
2	Evaluation questionnaire	Evaluation questionnaire		Evaluation questionnaire	CME research team
	Focus groups			Telephone interview	

### PAR questionnaires

To address the first question, the response of Nova Scotia physicians, patients, colleagues and coworkers to the Alberta instruments, Dalhousie CME recruited eligible NS physicians. Once recruitment was completed, the volunteer database was sent to CIS who managed the data collection, data entry, data analysis, and compilation of physician reports using the PAR instruments for the NS pilot study. In February 2002, CIS mailed the PAR packages to the physician volunteers. This included the self-assessment questionnaire, 25 patient questionnaires, and contact forms on which physicians were instructed to list the names of potential reviewers; i.e., 8 co-workers and 8 medical colleagues. Upon receipt of the names

and contact information for the co-workers and colleagues, CIS mailed questionnaires to these reviewers.

The PAR instruments and distribution methods are described below. For each questionnaire item on the self, medical colleague and coworker assessment questionnaires, reviewers were asked to rate themselves (for self) or the NSPAR participant (for colleagues and co-workers) compared to other physicians they know using a 5-point likert scale where a score of 1 is “among the worst”, 2 is “bottom half”, 3 is “average”, 4 is “top half” and “5” is “among the best”. The patient assessment questionnaire used a different 5-point likert scale, and asked respondents to rate their agreement with each statement about their physician, ranging from “strongly disagree” to “strongly agree”. Each PAR questionnaire also provided an option for the reviewer to indicate “unable to assess” for each item.

#### 1. Physician self-assessment questionnaire (SAQ)

Each physician received a questionnaire containing 31 statements about physician practices. For each statement, they were asked to rate themselves compared to other physicians they know.

#### 2. Medical colleague assessment questionnaire (MCAQ)

Each physician was asked to provide the names of 8 medical colleagues with whom they work or consult on a regular basis. Medical colleagues were defined as physicians who could make an informed response to the questions, for example, colleagues in their office, specialists, or consultants. CIS mailed a questionnaire to each medical colleague. The MCAQ was identical to the self-assessment questionnaire, and contained 31 statements about physician practices for five attributes: clinical competency, psychosocial management of patients, patient interaction, professional self-management, and consultation communication. Colleagues were asked to rate the physician compared to other physicians they know.

#### 3. Coworker assessment questionnaire (CWAQ)

Each physician was requested to provide the names of 8 of their non-physician coworkers. These include nurses, pharmacists, physiotherapists, office personnel, lab and x-ray technicians, and other health professionals with whom they work or share patients. As for the medical colleagues, CIS mailed a questionnaire to each coworker. This questionnaire contained 17 statements about physician practices for three attributes: patient interaction, coworker collegiality and co-worker communication. Co-workers were asked to rate the physician compared to other physicians they know.

#### 4. Patient assessment questionnaire (PAQ)

Each physician was requested to randomly select 25 patients to complete a PAR questionnaire. Physicians and/or their office personnel (i.e., receptionist or nurse) were instructed to provide each of the 25 patients with the PAR materials after their consultation with the physician. These materials included a cover letter explaining the project, the patient assessment questionnaire and an unsealed NSPAR envelope. Patients were asked to seal the completed questionnaire in the envelope provided and return it to the receptionist prior to leaving the office. The PAQ contained 40 statements about physicians’ practices for 7 attributes: patient interaction, phone communication, information for patients, personal communication, office staff, physical office and appointments. Patients were asked to indicate their level of agreement with each statement about their doctors’ behaviours.

#### *Dalhousie evaluation data collection tools*

To address the second research question, concerning the feasibility of using the PAR instruments and processes in Nova Scotia, all participants and reviewers were requested to complete an evaluation questionnaire developed by Dalhousie CME. Additionally, NSPAR

physician participants were invited to participate in a focus group and patients were invited to participate in a telephone interview.

The medical colleague, co-worker and patient questions were developed based on findings from the literature review and the Alberta PAR program. The physician evaluation questions were developed based on preliminary findings of the NSPAR reviewer questionnaires and the Alberta PAR program. All questionnaires were reviewed and discussed by the NSPAR research team and modifications were made based on their feedback.

The evaluation tools and data collection methods for each are described below.

#### 1. Physician evaluation questionnaire

The evaluation questionnaire consisted of 25 questions and took approximately 20 minutes to complete (see Appendix A for the physician evaluation questionnaire). It covered the following topics:

1. Recruiting reviewers and completing the questionnaires (i.e., pertinence of questions; attitudes about including the reviewers; comfort levels and difficulties recruiting reviewers)
2. Reviewing the NSPAR confidential report (i.e., interpreting the report; amount of time spent reviewing the report, and agreement with reviewers)
3. Demographic information
4. Factors that motivated the physician to participate in the NSPAR pilot project

Throughout the questionnaire there were places for comments and for respondents to specify details about certain responses.

The evaluation questionnaire was mailed to all NSPAR participants with their final report. Participants were asked to complete the questionnaire after reviewing their report, and then to return it to Dalhousie CME. One reminder letter was faxed to participants approximately 2 weeks after the report was sent.

#### 2. Medical colleagues and co-worker evaluation questionnaire

Medical colleagues and co-workers received identical evaluation questionnaires. The questionnaire consisted of 8 questions and took 5-10 minutes to complete (see Appendix B for the Medical colleague/co-worker evaluation questionnaire). It covered the following topics:

1. Reviewer's familiarity with the physician
2. Questions that were difficult to answer, and explanations of why they were difficult.
3. Attitudes about participating in the process (i.e., ability to answer truthfully; comfort level, and usefulness of the process)

There was also a general comments section at the end of the questionnaire for respondents to add additional information about participating in the NSPAR pilot process.

The evaluation questionnaire was mailed to all medical colleagues and co-workers with the PAR questionnaire. Reviewers were requested to complete the Dalhousie evaluation questionnaire after they completed the PAR questionnaire, then return it in a self-addressed, stamped envelope to Dalhousie CME.

### 3. Patient evaluation questionnaire

The patient evaluation questionnaire consisted of 11 questions and took approximately 10 minutes to complete (see Appendix C for the patient evaluation questionnaire). It covered the following topics:

1. Location where the questionnaire was completed
2. Ease and comfort of completing the questionnaire
3. Patient's familiarity with the physician
4. General comments about the questionnaire and the process
5. Demographic information

Patients were given both the PAR and Dalhousie evaluation questionnaires while in their physician's office. Although patients were requested to complete the PAR questionnaire immediately following their doctor's appointment, they were asked to complete the Dalhousie evaluation questionnaire at their home, and then to return it in a self-addressed stamped envelope to Dalhousie CME.

### 4. Patient interview

Question development:

The interview questions were developed based on preliminary findings of the Dalhousie evaluation questionnaires. The questions were reviewed by the NSPAR research team and changes were made based on their feedback. The interview questions consisted of 30 close-ended questions designed to explore patients' participation in, understanding of, and attitudes about the NSPAR process (see Appendix D for the patient interview questions).

Participant recruitment:

An invitation to participate in a 10-15 minute telephone interview was given to each patient with the Dalhousie evaluation questionnaire. Those who were interested in participating were asked to return an interview agreement form to Dalhousie CME.

Interview procedure:

Dalhousie CME hired and trained an interviewer specifically for the purpose of conducting the patient telephone interviews. The interviewer tried calling each of the 953 patient volunteers; however, due to time and resource constraints, only two phone attempts were made for each patient volunteer.

Each interview began with a standardized introduction. Participants were thanked for their participation. Then, the purpose and objectives of the interview were explained and participants were informed that their responses were confidential and they could choose to stop participating at any time during the interview. Participants were asked all 30 questions, in sequential order, and their responses were recorded directly into an ACCESS database during the interview.

### Physician participant focus group

Question development:

The focus group questions were developed based on preliminary findings of the PAR and Dalhousie questionnaires and suggestions from the CPSNS. A draft version of the questions was reviewed and discussed by the NSPAR research team. They thoroughly reviewed the initial questions and suggested changes, ensuring questions were appropriately addressing the two research questions (see Appendix E for the focus group questions).

Focus group questions were designed so that general open-ended questions were asked first, to facilitate discussion from all participants (i.e., reflecting back upon the NSPAR process, was NSPAR what you expected?), then questions became progressively more focused on the topics of interest.

**Participant recruitment:**

At the completion of the NSPAR process, participants were mailed their final NSPAR report summarizing results from their colleagues, coworkers, patients and self-ratings. Along with this report, participants were sent a focus group invitation letter. All participants were invited to take part in a 1.5 hour long focus group to discuss the acceptability and feasibility of the PAR process in Nova Scotia. Volunteers who could not attend the focus groups were invited to provide comments to the research team in writing or over the phone.

**Focus group procedure:**

Participants were faxed the questions prior to the focus group.

The Dalhousie CME Director of Program Development & Evaluation and the CME Research Associate facilitated all three focus groups. One facilitator asked the series of questions, ensured that all participants partook in the discussion, and summarized thoughts/key points at the end of each question. The second facilitator took notes throughout the discussions, and helped summarize key points and answer questions.

Each focus group began with a standardized introduction. Participants were welcomed and thanked for their participation. The purpose and objectives of the focus group were explained. Then, the focus group “ground rules” were explained, participants were informed that the focus group was being audio-taped and that their responses were confidential.

## **Analysis**

### *PAR questionnaire data*

CIS managed the data collection, data entry, data analysis, and compilation of reports using the PAR instruments for the NS pilot study. The unaggregated data for each assessment questionnaire (i.e., self assessment, medical colleague, co-worker and patient) were sent to Dalhousie CME upon completion of the pilot. All questionnaire data were exported to SPSS for statistical analysis. Frequency distributions and descriptive statistics (means and standard deviations) were computed for each question and each attribute of the PAR questionnaires.

Reliability coefficients for each scale and attribute were calculated using Cronbach’s alpha. For this purpose, missing data were replaced with the mean for each item. However, if the amount of data missing for a variable exceeded 30%, these items were removed and the alpha was calculated based on the remaining items.

Mean questionnaire scores were calculated for each PAR questionnaire (SAQ, MCAQ, CWAQ and PAQ) and mean attribute scores were calculated for each attribute of the MCAQ, CWAQ and PAQ (an attribute is composed of several items related to the same topic; e.g., “clinical competency” or “patient interaction”).

Inferential statistics (i.e., analysis of variance) were calculated to determine if there were any effects of reviewers’ familiarity with or relationship to the NSPAR volunteer on their mean

questionnaire scores and attribute scores. Independent sample t-tests were used to compare ratings on individual questionnaire items of the SAQ and MCAQ.

Measures of association (i.e., Pearson correlations) were calculated to determine if there were any relationships between ratings on the MCAQ and physicians' responses to the Dalhousie evaluation questionnaire, specifically with their opinions about recruiting reviewers, completing the questionnaires, and their agreement with the results. Only correlations significant at the  $p < .01$  level are reported.

#### *Dalhousie evaluation questionnaire data*

##### Quantitative data

Questionnaire data were entered into an ACCESS database, and then exported to SPSS for statistical analysis. Descriptive statistics (means and standard deviations) were calculated for data collected in Likert scales, and frequency distributions were calculated for non-continuous data.

##### Qualitative data

Qualitative data from the participating physician, medical colleague, co-worker and patient questionnaires were entered into an ACCESS database, coded and subjected to content analysis using a category system to identify themes and key concepts. Frequency distributions were calculated on coded data to determine how often certain themes were mentioned.

#### *Telephone interviews*

Interview data were entered into an ACCESS database, and then exported to SPSS for statistical analysis. Descriptive statistics (means and standard deviations) were calculated for data collected in Likert scales, and frequency distributions were calculated for non-continuous data.

#### *Focus groups*

All focus group sessions were audiotaped and transcribed. Transcriptions were coded by a researcher skilled in qualitative analysis; that is, they were broken down into meaningful pieces related to emerging themes. Transcripts and notes were reviewed several times to ensure that all relevant data were accounted for and systematically coded under the appropriate themes. Common themes and key concepts were identified and grouped into categories for each of the 3 focus groups. A second researcher reviewed the coding structure and codes assigned for the longest transcript, to ensure reliability of the process.

## RESULTS

### Response rates

#### *PAR process and Dalhousie evaluations*

Interest in the study far exceeded expectations. The initial goal was to recruit 100 NS family physicians. Due to the great interest in the project, this was increased to a maximum of 170. Physicians (n=2) who volunteered after the 170 were accepted were not able to participate.

Eight of the original 170 volunteers withdrew from the study: three because they were not in general office practice and five for other reasons including illness, unexpected travel and course work, unforeseen complications, NSPAR involving more work than expected, and other additional work demands. Of the remaining 162 volunteers, 142 completed the NSPAR process and received a final report. Table 2 shows the response rates of physicians, and their colleagues, coworkers and patients to the PAR and Dalhousie questionnaires.

**Table 2:** Response rates for NSPAR pilot project.

QUESTIONNAIRE	EXPECTED SAMPLE SIZE	ACTUAL SAMPLE SIZE BASED ON FPS PARTICIPATION	PAR QUESTIONNAIRE RESPONSES N (% ACTUAL SAMPLE)	DAL EVALUATION QUESTIONNAIRE RESPONSES N (% ACTUAL SAMPLE)
PHYSICIANS	167	142 (87.7%)	138 (97.2%)	113 (79.6%)
COLLEAGUES	1336	1136	930 (81.9%)	855 (75.3%)
		<i>Peer</i>	340 (36.6%)	396 (46.3%)
		<i>Consultant</i>	325 (34.9%)	459 (53.7%)
		<i>Referring physician</i>	73 (7.8%)	----
		<i>No response</i>	192 (20.6%)	----
COWORKERS	1336	1136	964 (84.9%)	959 (84.4%)
PATIENTS	4175	3550	3442 (97%)	2223 (62.6%)
PATIENT INTERVIEWS	1000		----	591 (59.1%)

#### *Demographic information for physician participants and patients*

##### Physicians

One hundred and forty-two physicians completed the NSPAR pilot project. Table 3 presents demographic information for these 142, the 113 who completed the Dalhousie evaluation questionnaire, and demographic information from the Medical Society of NS (2001) for NS family physicians. Of the 142 physicians completing NSPAR, 71.1% were male and 28.9% practiced in communities larger than 50,000 (i.e. Halifax Regional Municipality and Sydney), 34.5% in communities of 5,000-50,000, and 36.6% in communities smaller than 5,000.

Demographic information on the Dalhousie evaluation questionnaire revealed that 79% of respondents graduated from medical school between 1970-1989, 80% were between 41-60 years of age, 61% were in group practice and 38% were in solo practice. Respondents reported that they see an average of 148.5 (SD=58.0) patients per week.

Comparisons of this information to the Medical Society of Nova Scotia (2001) family and general practitioners' demographic information revealed that there were a larger percentage of men (71% vs 63.5%) and fewer metro physicians (28.9% vs 45.4%) in the NSPAR pilot project than is in the Nova Scotia physician population.

**Table 3:** Demographic information of three physician populations: NSPAR volunteers who participated in the pilot study, NSPAR volunteers who completed the Dalhousie evaluation questionnaire, and NS family physicians.

	<b>NSPAR participants</b> (N=142)	<b>Dal. evaluation questionnaire</b> (N=113)	<b>NS physicians</b> (N=903)
<b>Gender</b>			
Male	101 (71.1%)	84 (74.3%)	573 (63.5%)
Female	41 (28.9%)	29 (25.7%)	330 (36.5%)
<b>Community size</b>			
> 50,000	41 (28.9%)	32 (28.3%)	Metro: 410 (45.4%)
5,000 – 50,000	49 (34.5%)	50 (44.2%)	
< 5,000	52 (36.6%)	30 (26.5%)	Nonmetro: 493 (54.6%)

On the physician evaluation questionnaire, physicians were asked to specify the reason(s) why they participated in the NSPAR pilot project. As shown in Table 4, the main reason for volunteering was for self-improvement and/or self-exploration.

**Table 4:** Physicians' reasons for participating in NSPAR as reported on the physician evaluation questionnaire (note: physicians could indicate more than one reason)

<b>Reason for participating</b>	<b>Frequency</b> (n=103)
Self-improvement and/or self-exploration	95
Interest in the project	78
Practice improvement	70
5-year exemption from APMPR or other assessment program(s)	70
CME credit	41
An opportunity to influence the CPSNS' assessment program	39
Other	9

### Patients

In total, 3443 patients participated in the NSPAR pilot project. The majority of these patients (84.7%) reported that they had visited their doctor over three times in the last five years. Patients were asked to indicate the reason for their doctor's visit on the day they completed the PAR questionnaire. Results revealed that 54.5% were seeing their doctor for an ongoing

concern, 21.5% for a new concern, and 15% were there for an examination. An additional 9.1% did not specify the reason for their visit.

As shown in Table 5 below, the PAR patient sample was comprised of more women than men, more people who resided in a rural area than a town or city, and had a median age between 45-54 years of age. Compared to the Nova Scotia population demographics reported by Statistics Canada, 2001, this patient sample was slightly older (Nova Scotia median age is 38.8), and had a larger representation of women (NS stats: 48.4% males and 51.6% females). Statistics were not available on the distribution of the NS population by size of community, but it appears, from the table below, that there may have been over-representation of the rural population in the NSPAR study.

Patient respondents to the Dalhousie evaluation questionnaire provided information about their education level. Approximately 60% of this sample received formal education after high school. Specifically, 15.5% completed a university degree, 28.1% completed a diploma or certificate program, 15.8% had some post high school education, 20.6% completed high school, 15.2% had less than high school, and 4.7% did not provide this information.

**Table 5:** Demographic information of three patient study populations: Patients responding to the PAR questionnaire (N=3442), the Dalhousie CME questionnaire (N =2223), and participating in the interview (N=591).

	PAR responses N (%)	Dal responses N (%)	Patient interview N (%)
<b>Gender</b>			
Male	847 (24.6%)	648 (29.1%)	161 (27.3%)
Female	1977 (57.4%)	1476 (66.4%)	428 (72.7%)
No response	618 (18%)	99 (4.5%)	-----
<b>Age</b>			
25 or under	201 (5.8%)	95 (4.3%)	5 (.8%)
26-34	360 (10.5%)	168 (7.6%)	30 (5.1%)
35-44	695 (20.2%)	435 (19.5%)	97 (16.4%)
45-54	745 (21.6%)	505 (22.7%)	152 (25.7%)
55-64	583 (16.9%)	431 (19.4%)	143 (24.2%)
65 or over	625 (18.2%)	496 (22.3%)	164 (27.7%)
No response	233 (6.8%)	93 (4.2%)	-----
<b>Community size</b>			
Rural area or village	-----	926 (41.7%)	281 (47.5%)
Town	-----	521 (23.4%)	157 (26.6%)
City	-----	674 (30.3%)	152 (25.7%)

#### Focus groups

Thirty-nine NSPAR participants volunteered for the focus groups; four of these volunteered after the focus groups were completed. Due to scheduling difficulties many of the volunteers were

unable to attend, so they were invited to provide comments to the focus group questions in writing or over the phone to the research team. Two volunteers wrote notes and 4 phoned to give brief comments.

A total of 15 NSPAR volunteers participated in 3 focus groups, 5 females and 10 males. Three different settings were used to accommodate the volunteers: a teleconference (n=2), a face-to-face meeting during a CME event (n=9), and a combination face-to-face meeting and videoconference (n=4). Two of the focus groups were 1.5 hours long and one, the face-to-face group, was one hour long, due to time constraints. Representation from towns (population between 5,000 – 50,000) was lower than from urban (population > 50,000) and rural (population < 5,000) areas. In total, there were 8 rural, 2 town, and 5 urban participants.

### Research question #1

Research question #1 was “How do Nova Scotia physicians, their medical colleagues, coworkers and patients respond to a set of instruments developed and standardized in Alberta to assess physician performance?”

Each NSPAR volunteer was requested to recruit or provide names of 8 medical colleagues, 8 coworkers and 25 patients. The majority of participants fulfilled these PAR requirements, but in some cases, reviewers did not respond to reminders sent by CIS to complete the PAR questionnaires. As Table 6 shows, about 95% of physician participants provided names of eight medical colleagues and coworkers, while 87.3% of participants received responses from six or more medical colleague respondents. For coworkers, 94.4% of participants provided eight names and 94.4% also received responses from six or more. For patients, 98.6% of physician participants received responses from 18 or more patients.

**Table 6:** Number (%) of participants providing requested number of names of reviewers, receiving required number of responses, and mean, range of respondents per participant, by reviewer group. (N=142)

Type of Reviewer	# (%) of participants providing requested number of names	# (%) of participants with required number of respondents	# of respondents per participant		
			Mean	Min	Max
Medical Colleagues	136 (95.8%)	124 (87.3%)	6.6	0	8
Co-workers	134 (94.4%)	134 (94.4%)	6.8	0	8
Patients	N/A *	140 (98.6%)	24.2	0	26

\* Names not submitted; questionnaires distributed in office

To receive a complete NSPAR report (refer to Appendix J for a sample report), participants were required to have a minimum of 6 medical colleagues, 6 coworkers and 18 patients. If insufficient responses (i.e., < 60%) were received for individual questionnaire items, overall attributes or entire questionnaires, participants did not receive scores for that particular section. The report presents each physician with his or her mean scores for each item compared to the mean and range for the study group. They also receive a score for each attribute. (An attribute is composed of several items related to the same topic; e.g., “clinical competency” or “patient interaction”.) The report also uses “flags” as a type of recognition of performance.

Commendation flags are assigned to physicians receiving scores in the 90<sup>th</sup> percentile for items and attributes. Information flags are assigned whenever a physician's score is less than 4.0 and equal to or less than the 10<sup>th</sup> percentile.

One hundred forty-two physician participants received NSPAR reports. Table 7 presents overall results for each questionnaire. These results are discussed in more detail in the sections following this table.

**Table 7:** Descriptive statistics and reliability coefficients of the 4 PAR questionnaires for 142 Nova Scotia physicians.

Type of questionnaire	No. of items	No. of respondents	Return rate (%)	Mean item rating (and SD)	Reliability coefficient
Self-assessment	31	138	97.2%	3.73 (.47)	.96
Patient assessment	40	3442	97%	4.47 (.48)	.96
Co-worker assessment	17	964	84.9%	4.42 (.54)	.93
Medical colleague assessment	31	930	81.9%	4.27 (.54)	.97

### 1. Participant self-assessment questionnaire (SAQ)

As described above, the SAQ is comprised of 31 items and assesses five attributes. Summary descriptive data for the attributes of the 138 physician participants completing the SAQ are included in Table 8. The descriptive data for each questionnaire item, categorized by attribute, are reported in Appendix F.

**Table 8:** Descriptive data for the PAR attributes measured by the SAQ

Attribute	Number of reviewers	Mean	SD	Min.	Max.
Clinical Competency	138	3.65	.54	2.83	5.00
Consultation Communication	138	3.80	.52	3.00	5.00
Patient Interaction	138	3.98	.57	3.00	5.00
Professional Self Management	138	3.66	.49	2.63	5.00
Psychosocial Management of Patients	138	3.46	.59	2.00	5.00

The attribute receiving the highest mean score was patient interaction and the lowest was psychosocial management of patients. The overall mean across all 31 items and all respondents was 3.73. From Appendix G, the minimum score was 3.28 for item #19, "I manage patients with complex psychosocial problems", and the maximum was 4.09 for three items, all of which are part of the patient interaction attribute: #1 "I communicate effectively with patients", # 23 "I maintain confidentiality of patients and their families" and #24 "I respect the rights of patients". The mean SD across all items was 0.47.

The mean for item # 29, "I manage personal stress" was 3.54, with 11 (7.9%) physicians rating themselves as "2" or below. The College of Physicians and Surgeons of Alberta Physician Performance Committee, the committee responsible for the PAR program, uses this item as an important indicator of physician performance, and reviews the profiles of all physicians receiving a "2" or below.

### 2. Medical colleague questionnaire (MCAQ)

The MCAQ is identical to the SAQ, using the same 31 questions and five attributes. This was completed by 930 medical colleagues: 36.6% of these were peers of the NSPAR volunteer, 34.9% were consultants, 7.8% were referring physicians and an additional 20.6% did not specify their professional relationship to the volunteer. Summary descriptive data of the MCAQ attributes are presented in Table 9. Please refer to Appendix G for the summary by questionnaire item.

**Table 9:** Descriptive data for the PAR attributes measured by medical colleagues

Attribute	Number of reviewers	Mean	SD	Min.	Max.
Clinical Competency	905	4.22	0.59	2.17	5.00
Consultation Communication	913	4.32	0.57	2.50	5.00
Patient Interaction	905	4.33	0.59	2.67	5.00
Professional Self Management	906	4.22	0.57	2.17	5.00
Psychosocial Management of Patients	784	4.18	0.65	2.50	5.00

The overall mean across all 31 items and all respondents was 4.27. The minimum mean was 4.05, for item #30, "Is aware of own shortcomings" and the maximum was 4.39 for #22, "Shows compassion for patients and their families." The mean SD across all items was .54.

### 3. Coworker questionnaire (CWAQ)

The CWAQ is comprised of 17 questions and assesses three attributes. Summary descriptive data are presented in Table 10. Please refer to Appendix H for the summary by each questionnaire item.

**Table 10:** Descriptive data for the PAR attributes measured by co-workers

Attribute	Number of reviewers	Mean	SD	Min.	Max.
Patient Interaction	955	4.48	0.57	2.00	5.00
Co-worker Collegiality	963	4.46	0.61	1.25	5.00
Co-worker Communication	958	4.31	0.63	1.00	5.00

The overall mean across all 17 items and all respondents was 4.42. The minimum mean was 4.10, for item # 4, "Writes prescriptions clearly" and the maximum was 4.66 for item # 16, "Maintains confidentiality of patients". The mean SD across all items was .54.

#### 4. Patient questionnaire (PAQ)

The PAQ is comprised of 40 questions and these questions assess seven attributes. Summary descriptive data by attribute are presented in Table 11, and by item, in Appendix I.

**Table 11:** Descriptive data for the PAR attributes measured by patients

Attribute	Number of reviewers	Mean	SD	Min.	Max.
Appointments	3423	4.10	0.79	1.00	5.00
Information for Patients	3420	4.44	0.55	1.00	5.00
Office Staff	3408	4.53	0.59	1.00	5.00
Patient Interaction	3441	4.63	0.52	1.00	5.00
Personal Communication	3388	4.53	0.57	1.00	5.00
Phone Communication	2956	3.87	1.02	1.00	5.00
Physical Office	3421	4.49	0.56	1.00	5.00

The overall mean across all 40 items and all respondents was 4.47. The minimum mean was 3.70, for item #21, "I am able to reach a doctor by telephone after office hours", and the maximum was 4.75 for #39, "I would go back to this doctor". The mean SD across all items was .48.

#### *Comparisons of SAQ and MCQ responses*

The 31-items on the SAQ and MCAQ were identical. Independent sample t-tests were conducted to determine if there were any differences in physicians' self assessment responses compared to those from their medical colleagues. The results revealed that the mean ratings from the medical colleagues were significantly higher on each of the 31 items than the physicians' mean self-ratings ( $p < .001$  for each item). This is consistent with findings from the Alberta PAR pilot, reported by Hall et al, 1999. As the items on the coworker and patient questionnaires differed from those on the SAQ, it was not possible to conduct comparisons.

#### *Internal consistency of PAR questionnaires*

The reliability coefficient for each questionnaire and each attribute within each questionnaire (or scale) were calculated using Cronbach's Alpha. Missing values were replaced by the mean. As shown in Table 12, all reliability coefficients except one are above 0.86 and the latter, for "appointments" is above 0.70. Reliability coefficients above 0.70 show that the items in the scale are working well together.

**Table 12:** Reliability coefficients for PAR questionnaires and attributes within the questionnaires

<b>Questionnaire/Attributes</b>	<b>Number of Items</b>	<b>Alpha coefficient</b>
<b>Physicians' self assessment</b>	<b>31</b>	<b>.9585</b>
<b>Medical Colleague (Overall scale)</b>	<b>31</b>	<b>.9685</b>
Clinical Competency	6	.9131
Psychosocial Management of Patients	4	.8864
Patient Interaction	6	.9108
Professional Self Management	8	.8681
Consultation Communication	7	.9118
<b>Co-worker (Overall scale)</b>	<b>17</b>	<b>.9345</b>
Patient Interaction	8	.9078
Co-worker Collegiality	4	.8602
Co-worker Communication	5	.7555
<b>Patients (Overall scale)</b>	<b>40</b>	<b>.9642</b>
Patient Interaction	10	.9464
Phone Communication	2	.8233
Information for Patients	8	.9027
Personal Communication	6	.8867
Office Staff	6	.9543
Physical Office	5	.8819
Appointments	3	.7335

## Research question # 2

The second research question was, "Is it feasible to use the PAR instruments and processes in Nova Scotia?" To answer this question, we explored the following:

1. Physician participants' and reviewers' overall opinions about NSPAR process
2. Items receiving "unable to assess" responses from reviewers
3. Influence of familiarity on medical colleague, coworker and patient ratings
4. Reviewers' reported comfort with the NSPAR process
5. Physician participant perceptions of reviewer groups and their results
6. Physician participant perception of the usefulness of the report

### 1. *Physician participants' and reviewers' overall opinions about NSPAR process*

Physicians were given an opportunity on the Dalhousie evaluation questionnaire to comment about the NSPAR process, recruitment of reviewers and the final report. In total, 81 comments were provided. While the majority addressed specific aspects of NSPAR, 14 were general in nature. These participants indicated that they enjoyed the process, had no difficulties, and felt this method of evaluation was effective.

All focus group participants interacted in discussions about the overall or relative value of the PAR program. Most physician participants across the three groups agreed that NSPAR was generally useful for providing physicians with an overview of their practice. Some had also

participated in the APMPR process or had colleagues in their practice who had recently undergone review by APMPR. One physician participant had undergone the extensive week-long practice assessment review conducted by the University of Manitoba. These physicians compared NSPAR to those programs. The groups identified the following strengths of NSPAR:

1. NSPAR directly involves patients, which is most important. The involvement of coworkers is also helpful.
2. NSPAR, because of involvement of patients, coworkers and colleagues provides a broader scope of one's practice than APMPR, generally conducted by one assessor. However, the two are very different assessment programs and measure different things. APMPR focuses more directly on assessing office practices and administration, and quality of medical records.
3. There was some agreement that physicians should be accountable and should be reviewed on an on-going basis. One physician reported being in practice for 20 years and this was his first review. Being assessed on a regular basis through a broad-spectrum program such as NSPAR could provide both accountability to the public and useful feedback to physicians.

They also made suggestions for improving NSPAR. These included: improve the specificity of results to increase the usefulness of the report as an educational tool, add an objective measure such as chart audit to the assessment process, and address the uncertainty about the value of medical colleague feedback. These are discussed in more detail in this section.

Most physician participants agreed that NSPAR was useful as a process for providing NS physicians with a general overview of their practice. Although a small minority expressed skepticism of the reasons for CPSNS' implementing NSPAR, believing it was to "weed out bad doctors", there was general support of its use as a screening process for identifying physicians who might need a more detailed review. With regard to this, physicians requested that the College be clear and open about the purpose of NSPAR, if it were to be implemented.

From the perspectives of the three reviewer groups, 85.9% of medical colleagues and 97.1% of coworkers indicated that the process was somewhat or very useful. In the interviews, 81.5% of patients indicated that their physician would use the results of the survey to improve the care he or she provides (note: 9.2% indicated that their doctor did not need to improve). Additionally, 91.2% of patients indicated that the opinions of patients could help physicians do a better job.

## *2. Items receiving "unable to assess" responses from reviewers*

Several physician participants expressed concern about reviewers' ability to answer some of the PAR questions. In focus groups and on the post-survey evaluation, physicians provided their perspectives on questions that were difficult for reviewers to answer. They felt that reviewers, especially medical colleagues, were unable to accurately assess particular items; for example, it is difficult for colleagues to know how they manage psychosocial problems, perform technical procedures or keep their medical records. These assessments are especially difficult to make for physicians in solo practices or in rural communities. Several physicians indicated that their communication with their colleagues, especially consultants, is often limited to referral letters, providing little information about most of the items included in the medical colleague questionnaire.

Due to these concerns, the use of the "unable to assess" response on the PAR assessment questionnaires was explored to determine the following: the frequency of this response choice;

which questions received this response most often, and which reviewer groups had the most difficulty assessing the PAR participants. The findings revealed that a greater percentage of medical colleagues used the ‘unable to assess’ response, on more items, than co-workers and patients. The items that could not be assessed by over 10% of medical colleagues are summarized by attribute and type of colleague (i.e., peer or consultant) in Table 13 below.

**Table 13:** Number (%) of “unable to assess” (UA) responses by over 10% of medical colleagues, by item, type of colleague and attribute

Questionnaire item by reviewer and attribute	PAR questionnaire N (%) of Unable to Assess		
	All medical colleagues (N=930)	Peers only (N=340)	Consultants only (N=325)
<b>Clinical competency:</b>			
Perform technical procedures skillfully	308 (33.1%)	58 (17.1%)	154 (47.4%)
<b>Consultation communication:</b>			
Handles transfer of care	96 (10.3%)	25 (7.4%)	47 (14.5%)
Communicates referral information to patients	211 (23.1%)	107 (31.5%)	62 (19.1%)
Provides understanding for patient care responsibility	98 (10.5%)	49 (14.4%)	34 (10.5%)
<b>Patient interaction:</b>			
Communicates effectively with patient’s family.	171 (18.4%)	36 (10.6%)	82 (25.2%)
Maintains confidentiality	178 (19.1%)	22 (6.5%)	104 (32%)
Respects patient’s rights	109 (11.7%)	13 (3.8%)	59 (18.2%)
<b>Professional Self Management:</b>			
Maintains quality medical records	274 (29.5%)	38 (11.2%)	158 (48.6%)
Involved with professional development	176 (18.9%)	21 (6.2%)	99 (30.5%)
Accepts responsibility	104 (11.2%)	19 (5.6%)	50 (15.4%)
Manages resources efficiently	161 (17.3%)	38 (11.2%)	81 (24.9%)
Manages stress	283 (30.4%)	45 (13.2%)	153 (47.1%)
Aware of shortcomings	195 (21.0%)	37 (10.9%)	101 (31.1%)
After hours call system	139 (14.9%)	12 (3.5%)	79 (24.3%)
<b>Psychosocial Management of Patients:</b>			
Recognizes psychological aspects of illness	137 (14.7%)	12 (3.5%)	77 (23.7%)
Makes appropriate use of community resources for psychological aspects of care	343 (36.9%)	59 (17.4%)	183 (56.3%)
Appropriate referral for psychological aspects of illness	378 (40.6%)	61 (17.9%)	207 (63.7%)
Manages patients with complex psychosocial problems	331 (35.6%)	53 (15.6%)	177 (54.5%)

Ten percent or more of respondents indicated they were “unable to assess” the physician for 18 of the 31 MCAQ items. For 8 of these, over 20% were unable to assess the physician. Items for which the highest percentage of medical colleagues were unable to assess (i.e., >30%) were three in the “psychosocial management of patients” attribute, resulting in 49 physician participants not receiving a score for this attribute. One item in the “clinical competency,” attribute, performing technical skills, also received over 30% unable to assess responses. Also of importance, consultant physicians responded “unable to assess” much more frequently than peers.

In the Dalhousie evaluation, medical colleagues identified items that were difficult to answer. The questions they had the most difficulty with were the same ones which larger percentages of medical colleagues were “unable to assess”, as reported above. Of note, 44% of medical colleagues indicated that they had difficulty answering questions, as compared to 24% for coworkers and 7.3% of patients interviewed. Medical colleagues reported that the main reason they were unable to respond to specific items was that the items were “not applicable to their relationship with the physician”. The next most common reason was that they “did not know the answer”.

As shown in Table 14, this problem of “unable to assess” responses was less prevalent with coworkers and patients. For six of the 17 CWAQ items, 10% or more of respondents indicated they were “unable to assess” and there were only 2 items on the CWAQ that more than 20% of respondents were unable to assess: these were ‘makes appropriate use of community resources for psychosocial aspects of illness’ (30.8% were unable to assess) and ‘available for consultation about mutual patients’ (20.4%). In the Dalhousie evaluation questionnaire, explanations for this difficulty were that these questions were either not applicable to their relationship with the physician, or they did not know the answer.

For the patient questionnaire, 10% or more of respondents indicated they were “unable to assess” the physician for 6 of the 40 items, ranging from a low of 12% to a high of 18.4%. However, these items appear to describe situations that the patients had not encountered with their physician, probably due to the nature of their medical problems or the reason for this specific visit. (Please refer to table below).

**Table 14:** Number (%) of “unable to assess” (UA) responses selected by over 10% of co-workers and patients, by item and attribute

Questionnaire item by reviewer and attribute	PAR questionnaire N (%) of Unable to Assess
<b>Co-worker (CAQ)</b>	<b>N=964</b>
<b>Patient Interaction:</b>	
Non-judgmental of patients and family	102 (10.6%)
Respects the rights of patients to make informed decisions	141 (14.6%)
<b>Co-worker Communication:</b>	
Writes prescriptions clearly	111 (11.5%)
Appropriate community resource use for illness’s psychological aspects	297 (30.8%)
Accepts responsibility	135 (14%)
Available consults for mutual patients.	196 (20.4%)
<b>Patients (PAQ)</b>	
<b>Information for patients:</b>	
Serious problems are followed up	414 (12%)
<b>Personal communication:</b>	
Problem prevention explained	526 (15.3%)
Medicine directions explained	442 (12.8%)
Medicine side effects explained	495 (14.4%)
<b>Phone communication:</b>	
Able to reach doctor after hours	635 (18.4%)
Doctor available in emergencies	426 (12.4%)

### 3. Influence of familiarity on medical colleagues, coworker and patient ratings

Reviewers were asked on the Dalhousie post-survey evaluation if they knew the physician well enough to answer most of the PAR questions. The majority of reviewers felt that they did. As shown in Table 15, over 95% of coworkers and patients indicated they knew the physician well enough compared to 87% of medical colleagues.

**Table 15:** Reviewers’ responses to whether or not they knew the physician well enough to answer most of the questions adequately.

	Colleague (n=855)	Coworker (n=959)	Patient (n=2223)	Patient interview (n=589)
<b>Yes</b>	746 (87.3%)	911 (95%)	2182 (98.2%)	573 (97.3%)
<b>No</b>	100 (11.7%)	41 (4.3%)	35 (1.6%)	10 (1.7%)
<b>No answer or unsure</b>	9 (1.1%)	7 (.7%)	6 (.3%)	6 (1%)

Colleagues and coworkers were also asked to indicate how well they know the physician professionally, using a 5-point scale ranging from ‘not at all’ to ‘very well’. As shown in Table 16, over 95% of colleagues and coworkers felt they knew the physician professionally at least

somewhat well. These results were further explored by comparing the 2 types of medical colleagues used as reviewers: peers and consultants. The results of an independent samples t-test revealed a significant difference between the means of these two groups,  $t(850)=6.66$ ,  $p<.001$ . Peers of physicians felt they knew the physician significantly better than consultants (means: 4.37 vs. 4.00, respectively). 84.4% of peers indicated they knew the physician well or very well compared to only 73.8% of consultants.

On the PAR assessment questionnaires, as opposed to the Dalhousie post-survey evaluations, medical colleagues and co-workers were asked to specify how well they knew the PAR participant (i.e., not at all, not well, somewhat, well, or very well). Table 16 summarizes these results. (Note: fairly large percentages did not respond to this question.)

**Table 16:** Medical colleagues' and coworkers' familiarity with physician participant

How well they know the NSPAR volunteer	Medical colleagues Frequency (%)	Coworkers Frequency (%)
Not well	11 (1.2%)	2 (.2%)
Somewhat	122 (13.1%)	111 (11.5%)
Well	333 (35.8%)	343 (35.6%)
Very well	262 (28.2%)	276 (28.6%)
No response	202 (21.7%)	232 (24.1%)

We wished to determine if reviewers' assessment ratings for the PAR participants differed depending on how well they knew the physician. To do this, we calculated an analysis of variance (ANOVA) for both the medical colleague and co-worker data using reviewers' mean assessment scores as the dependent variable. Table 17 shows the mean questionnaire scores assigned for each reviewer group increased as they knew the physician better.

**Table 17:** Medical colleagues and co-workers familiarity ratings and the mean questionnaire score given to the NSPAR participant.

Familiarity rating	Medical Colleagues		Coworkers	
	N	Mean (SD)*	N	Mean (SD)*
Not Well	11	3.87 (0.6)	2	4.03 (1.0)
Somewhat	122	3.90 (0.5)	111	4.06 (0.6)
Well	333	4.26 (0.5)	343	4.33 (0.5)
Very Well	262	4.41 (0.5)	276	4.63 (0.4)
No familiarity rating	202	4.33 (0.6)	232	4.48 (0.5)

\* Mean, SD of mean assessment scores for this group of reviewers

The ANOVAs revealed a significant effect of familiarity on overall scores for both the medical colleagues and co-workers ( $p<.001$ ). Scheffe post hoc analyses were performed to identify which group means were significantly different from one another. For both colleagues and co-workers, those who knew the PAR participant well or very well gave significantly higher scores than those who knew the participant somewhat well ( $p<.005$ ) or not well ( $p<.001$ ). Additionally, those who knew the participant very well gave higher scores than those who knew him/her well ( $p<.001$ ). These are important findings.

In response to their reports, physician participants in focus groups and on the post-survey evaluation raised questions about the potential influence of familiarity with medical colleagues upon the scores they assigned. They were concerned that lack of familiarity would result in

lower scores; that because colleagues did not know them well, they would rate them lower. They again gave the example of consultants who only know the physician through referral letters. On the other hand, they were also concerned that selecting colleague reviewers who knew them well would result in potentially biased higher ratings. Their concerns support the results of the above analysis of influence of familiarity upon PAR ratings.

#### 4. Reviewers' reported comfort with the NSPAR process

On the Dalhousie post-survey evaluations, most reviewers reported being comfortable participating in the NSPAR process. As shown in Table 18, medical colleagues reported less comfort participating than coworkers and patients. There was no significant difference in the responses of medical peers and medical colleagues to this question.

**Table 18:** Reviewers' self-reported levels of comfort in participating in NSPAR.

	<b>Colleague (n=855)</b>	<b>Coworker (n=959)</b>	<b>Patient questionnaire (n=2223)</b>	<b>Patient interview (n=590)</b>
Very comfortable	515 (60.2%)	733 (76.4%)	1812 (81.5%)	382 (64.6%)
Somewhat comfortable	249 (29.1%)	184 (19.2%)	277 (12.5%)	197 (33.5%)
Somewhat uncomfortable	63 (7.4%)	29 (3%)	40 (1.8%)	9 (1.5%)
Very uncomfortable	16 (1.9%)	3 (.3%)	1 (0%)	1 (.2%)
No answer	12 (1.4%)	10 (1%)	93 (4.2%)	0

As shown in Table 19, the majority of reviewers felt that it was easy for them to answer the questions about the NSPAR volunteers in a truthful way. Again, the frequency of medical colleagues was lower than coworkers and patients. In the interviews, 99.5% of patients indicated that they could be completely truthful when answering the questions. Only 1.9% thought their doctor might see their answers and 32.2% indicated that it didn't matter.

**Table 19:** Reviewers' self-reported ability to answer the NSPAR questions truthfully.

	<b>Colleague (n=855)</b>	<b>Coworker (n=959)</b>	<b>Patient (n=2223)</b>
Very easy	377 (44.1%)	595 (62%)	1731 (77.9%)
Easy	397 (46.4%)	336 (35%)	459 (20.6%)
Difficult	61 (7.1%)	18 (1.9%)	29 (1.3%)
Very difficult	4 (.5%)	0	0
No answer	16 (1.9%)	10 (1%)	4 (.2%)

#### 5. Physician participant perceptions of reviewer groups and their results

The Dalhousie post-survey physician evaluation questionnaire collected data about physician participants' perceptions, for each reviewer group (colleagues, coworkers, patients), about the appropriateness of inclusion of the group in NSPAR, difficulty in recruiting, comfort in including, and degree to which participants agreed with the results provided. Table 20 presents these results by reviewer group.

**Table 20:** Percentage of physician participants agreeing with the following descriptors about the NSPAR recruitment process and results, for medical colleagues, coworkers, and patients

	<b>Colleagues (N=855)</b>	<b>Co-workers (N=959)</b>	<b>Patients (N=2223)</b>
Appropriate to include	91.1%	88.5%	94.7%
Had difficulty recruiting	17.7%	20.4%	5.3%
Felt comfortable recruiting	78.8%	80.5%	89.4%
Agreed with results	65.5%	79.6%	90.2%

Among the three groups, participants reported the highest agreement and comfort with recruiting patients, and the least difficulty in recruiting patients. Participants experienced the most difficulty recruiting coworkers (20.4%) as compared to the other two groups, but were generally consistent in their ratings of appropriateness and comfort in recruiting, and agreement with results of coworkers. Participant responses to the involvement of medical colleagues were more varied. Although 91% agreed that it was appropriate to include medical colleagues, 17.7% had difficulty recruiting them, while 78.8% felt comfortable recruiting them, and only 65.5% agreed with their results. Additionally, 9.7% of participants did not respond to the question concerning “agreement with medical colleague results”, compared with non-respondent rates of 0.9% - 2.7% across all other variables and reviewer groups.

Of the 113 physician participants responding to the Dalhousie evaluation, 18 made comments about difficulties or concerns with recruiting reviewers, especially colleagues, but for a few, also coworkers. Comments included discomfort due to imposing upon reviewers who were already very busy professionals, concern that some specialists were being asked to complete multiple reviews, difficulty recruiting, recruited reviewers not submitting the questionnaires, difficulties of recruiting in rural areas.

The 15 focus group respondents expressed agreement with the inclusion of patients as reviewers, and, believed overall, that NSPAR generally provides a broad perspective on their practice through the inclusion of multiple reviewers. On the other hand, they raised a number of questions and concerns about the inclusion of medical colleagues, due to their degree of familiarity with the physician being assessed. As discussed in earlier sections, they questioned the validity of the results from medical colleagues who may not be familiar with their practices, while wondering if inviting colleagues they knew well would result in overly positive scores. There also appeared to be general agreement that discrepancies existed between their patient and medical colleague feedback. While patients generally tended to provide the highest results, which were gratefully received, medical colleagues provided the lowest ratings, which could be troubling for the physicians.

Focus group participants explored the lack of physician experiences with the process of peer evaluation as a factor contributing to concerns about medical colleagues scores. It was suggested that physicians have not been taught how to evaluate their medical colleagues, and are not accustomed to receiving constructive criticism from them. One participant who had been a reviewer as well as being reviewed stated that judging others “is very tough to do” and other physician participants agreed.

Additional analyses

As reported above, findings from both the Dalhousie evaluation questionnaires and focus groups indicated physician participants expressed greater disagreement with medical colleagues' results than with those from patients or coworkers, after reviewing their report. This was further explored to determine if physicians' responses were related to the ratings they received from their medical colleagues.

Measures of association revealed the following significant relationships, as illustrated in Table 21: (However, it is noted that the frequencies of negative and neutral responses were small compared to positive responses, especially for the latter two relationships; i.e., "appropriate to include" and "comfort identifying".)

- Physicians' agreement with their colleagues' ratings was positively correlated with their mean assessment ratings from their colleagues,  $r_{xy} = .48$ ,  $p < .001$ . Physicians who received high scores tended to agree with their colleagues results, whereas those who received lower scores tended to disagree with them.
- Physicians' opinions about the appropriateness of including colleagues to assess their practice was positively correlated with their mean PAR score from their colleagues,  $r_{xy} = .35$ ,  $p < .001$ . Physicians who gave positive ratings on the appropriateness to include colleagues tended to be those who received higher mean ratings from their colleagues.
- Physicians' comfort with identifying medical colleagues was positively correlated with their mean PAR score from their colleagues,  $r_{xy} = .27$ ,  $p = .005$ . This indicates that physicians who were most comfortable identifying their colleagues tended to be those who received higher mean ratings from their colleagues.

**Table 21:** Relationship between physicians' overall mean MCAQ scores from their medical colleagues and their agreement ratings on three of the Dal evaluation questionnaires (n=113).

Rating on evaluation questionnaire	Agree with results from colleague		Appropriate to include colleagues		Felt comfortable identifying colleagues	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
Disagree	3	3.82 (.19)	1	3.77 (0.0)	5	4.10 (.30)
Neutral	24	4.09 (.27)	5	4.04 (.27)	15	4.06 (.29)
Agree	68	4.35 (.25)	52	4.22 (.28)	60	4.30 (.26)
Strongly agree	5	4.46 (.29)	49	4.36 (.26)	28	4.33 (.32)

6. *Usefulness of NSPAR report to physician participants*

Data were collected about the usefulness of the report from the physician participants' evaluation questionnaire and from focus groups. Of the 113 evaluation respondents, 88.5% felt the report and guide were clear and easy to read, 91% felt the tables were easy to understand and 88.5% felt that the report provided useful information. They were also asked about the amount of time they spent reviewing their report and with whom, if anyone, they had or intended to review the results. The findings included:

- 84.1% of respondents spent between 10 and 60 minutes reviewing their report.
- 81.4% of respondents reviewed their report 2 or 3 times (twice: 38.9%; three times: 42.5%)

## *Report to CPSNS of a pilot study of PAR in NS*

- 69% indicated that they had reviewed, or intended to review, the results included in their report with others. These respondents indicated that they had or would review the report with: spouse/significant other (n=57), medical colleague (n=39), co-worker (n=28), office staff (n=10), CME staff (n=4), medical consultant (n=2), College of Physicians and Surgeons (n=1), medical student (n=1), friend (n=1).

The evaluation also asked physician participants if their NSPAR results might cause them to consider making changes to their practice. Sixty-one percent indicated that they had or would consider making changes to practice, while 10.6% of respondents did not answer this question. Respondents identified 75 intended changes to practice, which included improving:

- written communication, particularly prescription writing and maintenance of medical records (n=14)
- patient phone communication and availability, both during and after hours (n=11)
- patient wait times and their accessibility for appointments (n=10)
- communication with patients, including general communication skills (n=6), being less judgmental of patients and their families (n=3), providing information about treatment (n=3), and asking about prescription medications (n=2),
- referral information and communication with colleagues (n=6),
- communication and relationships with co-workers (n=5)
- physical office space (n=5)
- changing or reviewing their treatment or technical skills (n=5)
- psychosocial aspects of practice (n=3)
- increasing CME activity (n=2).

Additional items, each mentioned once, were: improving professional self management, increasing patient and office staff education, increase use of available resources, and improving deficiencies. Almost 75% of the intended changes addressed communication issues, most with patients but also with medical colleagues and coworkers. Patient accessibility issues (wait times and ease in getting an appointment) were listed by 10 respondents. Three respondents indicated that they needed more time to think about the report, and two others mentioned that they did not feel there was sufficient information provided for them to identify the specific changes needed.

Although 88.5% of evaluation respondents believed the report provided useful information, a small number of these respondents and the 15 focus groups participants shared concerns about its usefulness and made suggestions for improvement. Thirteen respondents expressed concerns with the report, which were echoed by the focus group participants. Some had trouble using the report and interpreting their results, and others did not agree with the results. Participants in all focus groups expressed concerns that the reporting of the results was vague and impersonal; i.e., that the score per item compared to the mean and range for the group, was not specific enough to be helpful in knowing how to improve one's practice. For example, they believed that receiving a low score on the item, "Selects diagnostic tests appropriately", did not provide adequate direction for how to improve. Both evaluation respondents and focus group participants suggested improvements, such as the opportunity for the reviewer to add specific comments or examples which could make the results more user-friendly and supportive of change, or to provide an explanatory comment if the score is extreme. Another suggestion made by focus group participants was combining the NSPAR results with more objective data; e.g., by including objective chart review of a random number of patients, as is used in APMPR, to provide concrete feedback for improvement.

A small number of evaluation respondents and many focus group participants suggested that improvements to the scale would be helpful. They found that the clustering of the scores in the upper half of the scale led to lack of clarity about the concept of “average”. Although “3” was indicated as being the average score on the written PAR scale, this was not communicated in the results as scores of “3” were consistently below the mean scores, for all items.

Another concern expressed across all focus groups was that some patient items created unrealistic expectations for the kind of service family physicians could provide. For example, “I am advised of results of tests or X-rays” may create the expectation that patients should be called to the office and informed of all results, not just the abnormal ones. Some factors were believed to be beyond the physician’s control, such as physical features of the office and its location.

### **Limitations of the study**

The study may be limited by the volunteer nature of the physician sample, as the volunteers may differ in some ways from the overall Nova Scotia physician population. Additionally, proportionately fewer urban physicians participated in the NSPAR pilot study than rural physicians. There is the possibility that urban physicians may have responded differently.

## **CONCLUSIONS AND DISCUSSION**

### **Research question #1:**

**How do a sample of Nova Scotia physicians, their medical colleagues, coworkers and patients respond to a set of instruments developed and standardized in Alberta to assess physician performance?**

This question was answered by reviewing the aggregate PAR data as collected and provided by CIS, for recruitment of reviewers and questionnaire scores, and by calculating reliability coefficients of scales.

#### *Recruitment*

Recruitment by physician participants of adequate numbers of reviewers in each of the three groups did not appear to be problematic, although only 87% had the required number of medical colleagues as compared to 94% for coworkers and 99% for patients. However, the high numbers of “unable to assess” responses for medical colleagues, discussed in the next section, may indicate difficulty in recruiting colleagues who knew the participants well.

#### *Responses*

The mean scores for all items on all questionnaires were above “3” (where 3 = “average”, 4 = “in top half”). Mean scores for the physician participant self-assessment questionnaire ranged from 3.28 – 4.09. Response rates of reviewer groups were high, and the means of the scores assigned by reviewers were generally high: for medical colleagues, 4.05 – 4.39; coworkers, 4.10 - 4.66; patients, 3.7 – 4.75. The means of only three of the 40 patient items were below 4.1: waiting times for appointments, availability of the doctor after hours, availability in emergencies. High ratings by reviewers have been reported in other peer and patient assessment programs (Hall, 1999; Lipner, 2002). Although, overall, the NSPAR mean scores appear lower than those for the Alberta pilot, no statistical comparisons have been calculated with the Alberta data.

A comparison of the scores for the self-assessment and medical colleague questionnaires demonstrated that physician participants rated themselves significantly lower than their medical colleagues. This also was a finding in the Alberta pilot study. (Hall, 1999, p.55)

#### *Reliability coefficients*

The reliability coefficients for all scales were high, indicating their high internal consistency; i.e., the items worked well together.

### **Research question #2:**

**Is it feasible to use the PAR instruments and processes in Nova Scotia?**

This question was answered by reviewing the following: overall opinion of NSPAR, reviewers’ responses to completing the PAR questionnaires, physician participants’ perceptions of reviewer groups, usefulness of the NSPAR report to physician participants,

*Overall opinion of NSPAR*

The vast majority (88.5%) of the 113 physician participants completing the written post-survey evaluation felt that the NSPAR report provided useful information. Moreover, the consensus of the 15 focus group participants was that NSPAR was useful as a process for providing NS physicians with a general overview of their practice. Although a small minority expressed skepticism of the reasons for CPSNS' implementing NSPAR, believing it was to "weed out bad doctors", there was general support of its use as a screening process for identifying physicians who might need a more detailed review. With this in mind, physicians requested that the College communicate clearly and openly with the membership about the purpose of NSPAR, if it were to be implemented.

Physician participants identified that the strengths of NSPAR include: most importantly, the inclusion of patients; the broad perspective which it provides on one's practice; and the inclusion of Nova Scotia physicians in a regular and ongoing evaluation and feedback process. Suggestions for improvement included improving the specificity of results to increase the usefulness of the report as an educational tool, and addressing the uncertainty about the value of medical colleague feedback. These are discussed in more detail in this section.

The three reviewer groups generally believed that the NSPAR process was helpful. While 91% of patients agreed that the opinions of patients could help physicians do a better job, 97% of coworkers and 86% of medical colleagues indicated that they believed the process was useful.

*Reviewers' responses to PAR assessment questionnaires*

On the PAR questionnaires, the response option, "unable to assess" was used relatively frequently, especially by the medical colleague group. For 18 of the 31 items in this questionnaire, 10% or more of respondents indicated they were "unable to assess" the physician, and for 8 of these items, 20% or more were "unable to assess". Items for which the highest percentage of medical colleagues were unable to assess (i.e., >30%) were three in the "psychosocial management of patients" attribute, resulting in 49 physician participants not receiving a score for this attribute. One item in the "clinical competency" attribute, performing technical skills, also received over 30% unable to assess responses. Consultant physicians responded "unable to assess" much more frequently than peers. These findings raise questions about the usefulness of including these items in the questionnaires, and/ or criteria for reviewers, especially consultant physicians. Of additional interest, results of the AB PAR did not show a similar result, but indicated that "unable to assess" was used in fewer than 10% of responses. (Hall, 1999).

Analysis of the medical colleague and coworker PAR results also showed that those who knew the physician participants "well" or "very well" gave significantly higher scores than those who knew them "somewhat well" or "not well". These results differ from those of the Alberta pilot study in which it was found that those who knew the physicians less well, assigned higher scores, but this was a very small percentage (Hall, 1999, p.55). This may be an important area for further study, and may contribute to decisions about inclusion of appropriate medical colleague reviewers.

Although the three reviewer groups responding to the Dalhousie evaluation were generally positive about completing the questionnaires, there were differences among groups. Slightly smaller percentages of medical colleagues, as compared to patients or coworkers, reported that they knew the physician well enough to answer the questions, were comfortable answering the

questions, and believed the NSPAR process was helpful for physicians. An important difference was that a large proportion of medical colleagues (44%) reported having difficulty answering specific questions, as compared to coworkers (24%) and patients (7%). The most common reason given for this was, "not applicable to relationship with the physician."

*Physician participant perceptions of reviewer groups and the results they provided*

Physician participants believed it was appropriate to include patients as reviewers in the assessment process, had little trouble recruiting them and about 90% agreed with their feedback. They believed slightly less strongly that it was appropriate to include coworkers, had more difficulty recruiting them and about 80% agreed with their results. While 91% of respondents agreed that medical colleagues should be included as reviewers, and 79% felt comfortable recruiting them, only 65.5% agreed with their results. Further analysis of the latter finding showed that, after reviewing their medical colleague results, physician participants who received high scores tended to agree with their results, and those who received lower scores tended to disagree with them. On a related note, an external review of the Alberta PAR program (Nichols, 2001) found that physicians believed that medical colleagues generally inflate their ratings of each other, although not necessarily intentionally. These findings raise further questions about the value of medical colleague feedback.

The 15 focus group participants, although a small proportion of the 142 physician participants, elaborated upon concerns regarding the feedback of medical colleagues feedback. They feared that colleagues lacked adequate information about their practices to make assessments in an informed manner. They also discussed physicians' general lack of experience with providing and receiving evaluation feedback with colleagues on their clinical performance, suggesting this was an infrequent experience for most. There was consensus that, while patients appeared to rate physicians the highest, medical colleagues rated them the lowest. The latter feedback appeared troubling for many, although some viewed it as an opportunity to explore ways to improve relationships with medical colleagues.

*Usefulness of NSPAR report to physician participants*

The majority of respondents to the written post-survey evaluation felt that the report provided useful information, and 61% indicated that they had or were planning to make practice changes in response. Of the 75 intended changes identified, almost 75% addressed communication issues, most with patients but also with medical colleagues and coworkers. The three most frequently identified intended changes were written communication, particularly prescription writing and maintenance of medical records (n=14); patient phone communication and availability, both during and after hours (n=11) and patient wait times and accessibility for appointments (n=10). In comparison with other studies, a follow-up study of the Alberta PAR pilot found that feedback from PAR led 83% of physicians to contemplate making a change and 66% to initiate a change. (Fidler, 1999). A recent study conducted by the American Board of Internal Medicine on their "peer and patient assessment module", a program similar to PAR, found that 65% of physicians believed that their feedback would help them improve the quality of their medical care. (Lipner, 2002).

The 15 focus group participants, and several evaluation respondents, offered a number of suggestions for improving the specificity of the report to increase its usefulness as a tool for continuing professional development. These included rewording some items, adding space for reviewer comments or examples, revising the scale, and adding an objective measure such as chart audit to the process.

## **Summary of key conclusions**

1. The vast majority (88.5%) of physician participant respondents to the written post-survey evaluation felt that the NSPAR report provided useful information, and 61% indicated that they had made or were planning to make practice changes in response to it. This is generally positive feedback on the usefulness of the NSPAR process. Of the 75 intended changes identified, almost 75% addressed communication issues, most with patients but some also with medical colleagues and coworkers.
2. Overall, most physician participants agreed that NSPAR was useful as a process for providing NS physicians with a general overview of their practice. Strengths of NSPAR are: the inclusion of patients, the broad perspective which it provides on one's practice, and the inclusion of Nova Scotia physicians in a regular and ongoing evaluation and feedback process. Areas for improvement include lack of specificity of results which diminishes its usefulness as an educational tool, and concerns that medical colleague feedback is of less value than that of other reviewers.
3. While the item means on the medical colleague, coworker and patient questionnaires were generally high (above 4.0 for all but 3 of the total 88 items), analysis of the results showed that medical colleagues and coworkers who knew the physician participants well or very well gave significantly higher scores than those who knew them somewhat well or not well. This important finding is worthy of further exploration.
4. Medical colleagues more frequently indicated they were "unable to assess" the physician than other reviewer groups, and consultant physicians responded "unable to assess" much more frequently than peers. The most common reason given for this was that the question was, "not applicable to my relationship with the physician".
5. A large proportion of medical colleagues (44%) reported having difficulty answering specific questions, as compared to coworkers (24%) and patients (7%).
6. While 91% of physician participants agreed that medical colleagues should be included as reviewers and 79% felt comfortable recruiting them, only 65.5% agreed with their results. Focus group respondents provided some explanations for this.
7. Focus group participants reported that receiving negative feedback; i.e., being rated below the mean or receiving an information flag, was troubling for them. This suggests a need for delivering the feedback in a sensitive and careful manner, as well as attempting to ensure the participation of reviewers who know physicians sufficiently well.
8. If NSPAR were to be accepted, physicians request clear and open communication by the College about the purpose of NSPAR and use of NSPAR data.

## **RECOMMENDATIONS**

If the CPSNS were to adopt the PAR program, we make the following recommendations:

### **Communication and support**

1. Communication with membership is fundamental to the success of new initiatives, as demonstrated by the College's successful NSPAR pre-pilot communication process. If NSPAR were accepted, we would again encourage clear, open and consistent communication with members about its rationale, purpose, processes and confidentiality procedures.
2. Although many respondents reported finding the report useful, some had difficulty interpreting the results and understanding how they might improve their practice. We would recommend the provision of a confidential counseling service to assist physicians with understanding and interpreting their reports, identifying ways to improve their practice, and accessing resources for continuing professional development.

### **Medical colleague reviewers**

3. As the medical colleague reviewer group appeared to respond differently than the other groups; i.e., appeared to be less familiar with the physicians, had more difficulty responding to items, and were unable to respond to more items than other groups, we suggest the following:
  - 3.1 Conduct further analyses of medical colleagues responses and results to determine if other factors influence their ability to respond and physician participants' responses to their results; e.g., geographic or practice type factors. This is important as few studies have been conducted in this area.
  - 3.2 Consider including a means of adjusting results to reflect level of familiarity with the physician.
  - 3.3 Consider excluding items which a larger percentage of medical colleague reviewers are "unable to assess", or assessing them in a different way; e.g., those relating to psycho-social care. The latter approach may be preferable. However, this changes the assessment tool and would require further study.
  - 3.4 As consultant medical colleagues reported more "unable to assess" items than peer colleagues, consider a means for identifying consultants who know physicians sufficiently well to complete the items on the questionnaire. It may also be helpful to review the separate consultant questionnaire used in the original Alberta PAR pilot study (Hall, 1999), and discuss its use with the Alberta research team.

Finally, regardless of the College's decision to adopt NSPAR, we recommend providing a summarized report of pilot study findings to NSPAR physician participants. They expressed much interest throughout the process and appear sincerely interested in its outcome.

## **REFERENCES**

Fidler H, Lockyer JM, Toews J, Violato C (1999). Changing physicians' practices: the effect of individual feedback. Acad Med; Jun;74(6): 702-14.

Hall W, Violato C, Lewkonja R, Lockyer J, Fidler H, Toews J, Jennett P, Donoff M, Moores D (1999). Assessment of physician performance in Alberta: the physician achievement review. CMAJ; 161(1):52-57.

Lipner RS, Blank LL, Leas BF, Fortna, GS (2002). The value of patient and peer ratings in recertification. Acad Med; Vol 77 (10 Suppl.): S64-66.

Nichols Applied Management (2001). Final report: evaluation of the Physician Achievement Review Program (PAR). Management and Economic Consultants; June, 2001.

Violato C, Marini A, Toews J, Lockyer J, Fidler H (1997). Assessing physicians in practice: feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. Acad Med; Oct; 72 (10 Suppl 1): S82-4.