

Assessment of Psychiatrists in Practice Through Multisource Feedback

Claudio Violato, PhD¹; Jocelyn M Lockyer, PhD²; Herta Fidler, MSc³

Objective: To assess the feasibility and evidence for the reliability and validity of a set of questionnaires for psychiatrists, given that multisource feedback (MSF) or 360° evaluation allows medical colleagues, coworkers, and patients to provide feedback about competencies to enhance physician improvement in intended directions.

Method: Surveys with 40, 22, 38, and 37 items were developed to assess psychiatrists by 25 patients, 8 coworkers, 8 psychiatrist colleagues, and self, respectively, using a 5-point agreement scale with an unable-to-assess category. Items addressed key competencies related to communication skills, professionalism, collegiality, and self-management. Feasibility was assessed with response rates for each instrument. Validity was assessed with a table of specifications, the percentage of participants unable to assess the psychiatrist for each item, and exploratory factor analyses to determine which items grouped together into scales. Reliability was assessed by Cronbach's alpha and generalizability coefficients.

Results: A sample of 101 psychiatrists provided data. A total of 2456 patients (24.32/25.00 per psychiatrist), 744 coworkers (7.37/8.00 per psychiatrist), 764 colleagues (7.56/8.00 per psychiatrist), and 101 self forms were analyzed. The overall internal consistency reliability of the instruments was a Cronbach's alpha of 0.98, 0.96, and 0.98 for patient, coworker, and medical colleague surveys, respectively. The generalizability coefficient for the patient, coworker, and medical colleague was 0.78, 0.82, and 0.81, respectively.

Conclusion: It is possible to develop a feasible MSF program for psychiatrists with evidence of reliability and validity that can provide feedback about key clinical competencies.

Can J Psychiatry 2008;53(8):525–533

Clinical Implications

- Patients, nonmedical coworkers (such as, psychologists and nurses), and medical colleagues will complete MSF questionnaires to provide feedback about observable behaviours.
- Data about aspects of collegiality, professionalism, communication skills, clinical skills, and professional development can be obtained through questionnaires.
- Questionnaires can be developed that are acceptable and provide evidence for validity and reliability.

Limitations

- The study was done under the auspices of a regulatory authority in one Canadian province and may have limited generalizability.
- Study participants selected their own respondents, which may have introduced a bias.
- It is not known how the psychiatrists used the data to change their practices.

Key Words: *psychiatrist, multisource feedback, 360-degree evaluation, physician assessment, professionalism, communication skills, professional development*

Medical organizations in Canada, the United States, and the United Kingdom have focused considerable attention on the core competencies required of physicians in practice. This has resulted in the creation of the CanMEDS,¹ Accreditation Council for Graduate Medical Education,² and Good Medical Practice³ competencies. Without reducing the centrality of medical knowledge and patient care expertise, these new competencies recognize the importance of professionalism, communication skills, team work, and professional development. They are used to guide the development and assessment of post-graduate trainees and inform the practices of licensed physicians. Tools such as direct observation, checklists, portfolios, global ratings, 360° evaluations, and OSCEs (Objective Structural Clinical Examinations)⁴⁻⁶ are advocated to assess trainees. It is more challenging to enhance these competencies for practising psychiatrists.

MSF or 360° feedback was advocated^{4,5} for the assessment of the broader competencies required for clinical practice. In medicine, MSF relies on aggregate data from questionnaires completed by patients, medical colleagues (such as peers and referral physicians), and coworkers (such as nurses, social workers, and psychologists) to provide feedback.⁶⁻¹⁰ Its purpose is to guide self-development by providing feedback about observable behaviours that can be addressed and possibly changed. This form of evaluation is currently used as part of revalidation and quality improvement for practising physicians in specialties such as family medicine,⁷ internal medicine,⁸ anesthesiology,⁹ and surgery.¹⁰

Studies of MSF show that reliable and valid instruments (questionnaires) can be developed.⁶⁻¹⁰ It appears feasible to develop quality improvement programs in which most of the physicians in the discipline can be assessed by 8 to 10 coworkers and 8 to 10 medical colleagues and 25 patients.⁶⁻¹⁰ This number of raters produces acceptable reliability for both the overall instrument and the physician being assessed.⁶⁻¹⁰ Further, given that the intent of MSF is to guide professional development, studies have shown that most participating physicians will use their feedback data to guide the changes they make.^{8,11} With feedback, physicians increased their explanations to patients, improved the print material in their offices, changed their communication strategies with their peers, and improved their psychosocial skills following feedback.¹¹

Abbreviations used in this article

CPSA-PAR	College of Physicians and Surgeons of Alberta—Physician Achievement Review
MSF	multisource feedback

The CPSA-PAR program began developing MSF instruments in 1996.^{7,12} Their program requires that every physician participate on a 5-year cycle. The program's original goal was to provide feedback to physicians about 6 broad categories of performance: medical knowledge and skills, attitudes and behaviour, professional responsibilities, practice improvement activities, administrative skills, and personal health.⁷ Instruments were developed and tested for most specialty groups.¹²

As part of the CPSA-PAR development work, one set of instruments was developed to be used for the combined medical specialties of pediatrics, internal medicine, and psychiatry.^{12,13} The instruments were developed by a working group that included psychiatrists along with specialist and subspecialists in internal medicine and pediatrics. To ensure that the instruments were appropriate for all of the specialty groups, every physician to be assessed with the set of instruments was provided an opportunity to review the questionnaires and provide feedback that was incorporated into the final set of instruments. Our earlier examination of the medical colleague instrument component demonstrated it was reliable and appropriate for use in all 3 disciplines.¹³ A comparison between the self and medical colleague data showed that most psychiatrists assessed themselves as average.¹⁴

Our main purpose was to extend the examination of the psychiatrist data specifically examining the patient, coworker (such as nurses and psychologists), and self data to assess the feasibility, validity, and reliability of an MSF system for psychiatry practice, in conjunction with the already reported data from the medical colleague assessment.^{12,13} Several questions were of interest:

- What is the feasibility of an assessment system for psychiatrists that provides feedback from patients, coworkers, medical colleagues, and self?
- What questions about psychiatrists' practices can patients, coworkers, and medical colleagues answer?
- What are the score profiles for each of the items (that is, mean and SD) on the surveys?
- Do the items on a survey group together into meaningful scales to guide performance improvement direction?
- Are the instruments reliable both for the practice of psychiatry and for the individual physician who is assessed?

Method

The psychiatrists were recruited by Pivotal Research Inc (formerly, Customer Information Services), a private company that handles the PAR work. With direction from the CPSA, Pivotal Research Inc recruited psychiatrists who were licensed to practice and had been in practice for 3 years. They

used the CPSA's list of licensed psychiatrists and drew a random sample from that list.

The final instrument for patients consisted of 40 items (Table 1). Raters were asked to use a 5-point rating scale (1 = strongly disagree, to 5 = strongly agree). The instrument for coworkers (Table 2) and medical colleagues (Table 3) consisted of 22 and 38 items, respectively, with the same 5-point rating scale. The self-assessment instrument and medical colleague questionnaires (Table 3) were identical except that for the self-assessment all items were written in the first person and the last item on the colleague questionnaire—"If a member of my family needed care, I would rate this physician"—was omitted. All questionnaires provided respondents with the option of indicating they were unable to assess the physician on the item.

Each participating psychiatrist was responsible for completing a self-assessment and identifying the 8 medical colleagues and 8 coworkers who could answer the questions on the survey. Pivotal Research Inc provided coworkers and medical colleagues with copies of the questionnaire and responses went directly to them.

Previous work established that raters chosen by people being assessed do not provide significantly different evaluations than those selected by a third party.¹⁵ Furthermore, studies examining how well the assessor and assessed physician knew one another showed that familiarity contributed very little to the variance in ratings.^{7,8} Each psychiatrist was provided with 25 copies of the patient survey. The physicians were provided with sealable envelopes so that the patients could complete the questionnaires anonymously in the physician's office and the physician's staff could send back all of the forms to Pivotal Research Inc for processing. Previous studies showed that 8 medical colleagues and coworkers and 25 patients would provide stable data for each participant (that is, a generalizability coefficient [Ep^2] ≥ 0.70).⁷⁻¹⁰

Numerous statistical analyses were undertaken to address the research questions posed. Response rates were used to determine feasibility for each of the respondent groups (question 1). For each item on each survey, the percentage of unable-to-assess items, along with the mean and SD, were computed to determine the viability of items and the score profiles (questions 2 and 3, respectively). When the percentage of unable-to-assess items exceeds 20% on a survey, it suggests a need to examine the item for revision or deletion. We used exploratory factor analysis to determine which items on the patient and coworker surveys belonged together (that is, became a factor or scale) (question 4). This analysis allowed us to identify the factors and numbers of factors for each instrument, and to describe the relative variance accounted for by each factor and their coherence. These factors or scales

could then be used to establish the key domains (for example, communication) for improvement while the items within each factor would provide more precise information about behaviours (for example, is courteous to coworkers). The factor analysis for the medical colleague instrument was previously reported.¹³ The factors for the self-assessment were not determined as the CPSA uses the factors from the medical colleague factor analysis to report self data. Last, reliability was assessed (question 5). Internal consistency reliability was examined using Cronbach's alpha coefficient for each of the rater groups and for each of the scales or factors for each rater group, which enables an assessment of overall instrument stability. This analysis was followed by a generalizability analysis to determine the generalizability coefficient to ensure there were sufficient numbers of items and raters to provide stable data for each individual psychiatrist on each instrument. Normally, $Ep^2 \geq 0.70$ suggests data are stable.^{6-10,15} If the generalizability coefficient is low, it suggests that more raters or more items would be required to enhance stability.

The study received approval from the University of Calgary, Conjoint Health Research Ethics Board.

Results

A total of 101 psychiatrists registered with the CPSA participated, producing 101 self-assessments (100% return rate). A total of 2456 (97.3%) patient surveys were available for a mean of 24.32 per psychiatrist. Nonphysician coworkers contributed 744 (92.1%) responses for a mean of 7.37 per psychiatrist. As previously noted, 764 medical colleagues (94.6% response rate) provided a mean of 7.56 assessments for each psychiatrist.¹³

Most items on the questionnaires could be answered by the respondents. As presented in Tables 1 to 3, 7/40 items on the patient survey, 0/22 items on the coworker survey, and 6/38 items on the medical colleague had unable-to-assess rates of more than 20%. The mean ratings for all items on the patient and medical colleague questionnaires were more than 4/5. There was one item on the coworker questionnaire (that is, writes legibly) with a mean rating of 3.91. There were several items on the self-assessment with mean ratings between 3 and 4.

The factor analysis identified 5 factors on the patient survey: technical communication, humanistic qualities, staffing, personal communication, and office structure, which accounted for 73.7% of the variance. There were 3 factors on the coworker questionnaire: humanistic and psychosocial, coworker collegiality, and written communication, which accounted for 68.8% of the total variance. As previously noted, there were 4 factors on the medical colleague survey that accounted for 66.8% of the variance: humanistic and

Table 1 Patient descriptive statistics, item analyses, and rotated component matrix for psychiatrists

Research item	<i>n</i>	Mean (SD)	UA, % ^a	TC ^b	H	S	PC	O
1. Explained my illness or concern to me clearly	2322	4.60 (0.66)	5.5				0.659	
2. Explained my treatment choices or options	2323	4.58 (0.67)	5.4				0.643	
3. Explained my follow-up plan to me	2279	4.54 (0.69)	7.2				0.680	
4. Told me how and when to take my medicine, if medicine was prescribed	2153	4.69 (0.60)	12.3				0.612	
5. Told me of side effects of the medicine, if medicine was prescribed	2132	4.47 (0.81)	13.2				0.685	
6. Spends enough time with me	2439	4.60 (0.69)	0.7		0.574			
7. Shows interest in my problems	2444	4.71 (0.59)	0.5		0.656			
8. Asks details about my personal life when appropriate	2420	4.66 (0.62)	1.5		0.538		0.403	
9. Answers my questions well	2428	4.66 (0.62)	1.1		0.610		0.441	
10. Examines me appropriately for my problems	2260	4.59 (0.65)	8.0		0.530		0.477	
11. Treats me with respect	2430	4.80 (0.49)	1.1		0.710			
12. Helps me with my fears and worries	2387	4.61 (0.66)	2.8		0.568		0.437	
13. Is easy to get into (for example, wheel chair accessible, parking)	2262	4.37 (0.84)	7.9					0.691
14. Has appropriate waiting areas	2374	4.55 (0.67)	3.3					0.773
15. Examining rooms are adequately sized and have adequate equipment	2033	4.56 (0.64)	17.2					0.757
16. Is clean and in good repair	2361	4.63 (0.58)	3.9					0.714
17. Provides adequate privacy	2365	4.63 (0.65)	3.7					0.736
18. I can reach the office by phone during the day	2326	4.49 (0.70)	5.3	0.440				
19. I receive an appropriate explanation if my appointment is delayed	2115	4.46 (0.76)	13.9	0.468		0.430		
20. My messages are returned	2184	4.55 (0.71)	11.1	0.521		0.457		
21. Staff: Are helpful and pleasant	2272	4.60 (0.63)	7.5			0.823		
22. Staff: Are respectful of patients	2268	4.61 (0.62)	7.7			0.858		
23. Staff: Behave in a professional manner	2265	4.63 (0.61)	7.8			0.839		
24. Staff: Work well with the doctor	2069	4.61 (0.63)	15.8			0.755		
25. Staff: Prevent patients from hearing confidential information about other patients	2086	4.55 (0.71)	15.1			0.716		
26. In an emergency situation, this doctor's office provides me with clear instructions on what I am to do	1818	4.42 (0.84)	26.0	0.661				
27. This doctor provides reports to my family doctor	1699	4.29 (0.92)	30.8	0.643				
28. When asked, this doctor provides insurance and medicolegal reports in a timely manner	1529	4.52 (0.72)	37.7	0.752				
29. When asked, this doctor provides reports, files, or copies of letters in a timely manner	1789	4.59 (0.65)	27.2	0.715				
30. This doctor arranges appointments with other specialists when necessary	1559	4.52 (0.72)	36.5	0.702				
31. This doctor's office follows up on serious problems	1935	4.60 (0.65)	21.2	0.659				
32. I am told what to do if my problems do not get better	2182	4.54 (0.71)	11.2	0.576				
33. I am asked about prescription and nonprescription medicine I may be taking	2329	4.56 (0.67)	5.2	0.459				

continued

Table 1 continued								
Research item	<i>n</i>	Mean (SD)	UA, % ^a	TC ^b	H	S	PC	O
34. This doctor talks to me about preventative care (for example, quitting smoking, weight control, sleeping, alcohol, exercise, etc)	2133	4.48 (0.74)	13.2	0.535				
35. This doctor has good written health information	1919	4.38 (0.77)	21.9	0.545				
36. This doctor refers me to appropriate educational resources (such as, websites, brochures patient support groups, books)	2067	4.33 (0.85)	15.8	0.548				
37. I would go back to this doctor	2424	4.80 (0.52)	1.3		0.807			
38. I would send a friend to this doctor	2385	4.75 (0.59)	2.9		0.743			
39. This doctor presents him or herself in a professional manner	2432	4.81 (0.48)	1.0		0.777			
40. I was helped by this doctor	2390	4.78 (0.53)	2.7		0.754			
Cronbach's α (overall = 0.98) ^c				0.94	0.90	0.95	0.79	0.83
Variance for each factor, % (total = 68.97)				51.60	6.40	4.37	3.60	3.00
H = humanistic; O = office; PC = personal communication; S = staff; TC = technical communication; UA = unable to access								
^a 20% indicated UA								
^b Extraction method: principal component analysis. \rightarrow Rotation method: Varimax with Kaiser normalization rotation converged in 7 iterations.								
^c Generalizability coefficient, $E_p^2 = 0.78$								

communication, psychosocial management of patients, clinical performance, and professional self-management.¹²

Cronbach's alpha was calculated to determine the internal consistency reliability of the instruments. There was an overall alpha of 0.98, 0.96, 0.98, and 0.96, respectively, on the patient, coworker, medical colleague, and self surveys. The generalizability analysis indicated that the generalizability coefficient values were 0.78 for patient surveys, 0.82 for coworker surveys, and 0.81 for medical colleagues.

Discussion

In this study, we evaluated a set of MSF data collected to assess psychiatrists' practices. The psychiatrists were assessed on numerous aspects of practice that the regulatory authority and the physicians themselves (through their participation on the committee and feedback about the questionnaires) believed to be important. While this quality improvement tool was not designed to specifically assess CanMEDS competencies, the items and the factors suggest a close alignment with these competencies.

The PAR program is mandatory and the response rates were high. As such, these rates are consistent with the response rates for other groups of physicians we have studied,^{7,9,10} and higher than achieved in US⁸ and UK⁶ studies.

While most of the items could be answered by all respondents, there were specific types of items on the medical colleague

and patient questionnaires that had unable-to-assess percentages higher than anticipated. For medical colleagues, these tended to be in aspects of professional management (such as, practising within scope of practice and stress management) or professional development, which colleagues may not observe. Among items that patients had more difficulty answering were ones that many patients would not have experienced in their relationship with their psychiatrist (seeking help in emergency situations, and writing reports and arranging appointments on patients' behalf). The decision to develop a set of questions that would work for psychiatry, and the specialties and subspecialties of internal medicine and pediatrics, may partly explain why not all questions were effective.

The score profiles are positively skewed. The range and the mean ratings were with most physicians receiving all of their ratings between 4 and 5. These profiles are similar to that of other groups of practising physicians^{7-10,15} and are consistent with the range of scores found in assessments of residents and medical students.⁶

Our exploratory factor analyses found that items did group together in meaningful and intended ways that are consistent with the intent of the PAR program. The CPSA, as a regulatory authority, is concerned about communication, humanistic qualities (that is, professionalism), clinical performance, and self-management (balance between personal and professional life). Items related to these attributes did group

Table 2 Coworker descriptive statistics, item analyses, and rotated component matrix for psychiatrists

Research item	<i>n</i>	Mean (SD)	UA, % ^a	Humanistic and psychosocial	Coworker collegiality	Written communication
1. Communicates effectively with patients	702	4.51 (0.69)	5.60	0.74		
2. Verbally communicate effectively with other professionals	734	4.48 (0.73)	1.30	0.64		
3. Effectively communicates in writing with other professionals	688	4.43 (0.74)	7.50	0.44		
4. Writes legibly	702	3.91 (1.03)	5.60			0.81
5. Is courteous to coworkers	726	4.63 (0.66)	2.40	0.61		
6. Demonstrates appropriate concern for worker safety	655	4.51 (0.70)	12.00		0.66	
7. Respects the professional knowledge and skills of coworkers	732	4.57 (0.72)	1.60		0.66	
8. Collaborates well with coworkers	725	4.48 (0.76)	2.60		0.66	
9. Shows compassion to patients and their families	712	4.58 (0.65)	4.30	0.79		
10. Separates personal values from the management of patients	662	4.45 (0.72)	11.00	0.73		
11. Is courteous to patients and their families	705	4.61 (0.62)	5.20	0.83		
12. Respects the rights of patients to make informed decisions	688	4.60 (0.66)	7.50	0.67		
13. Accepts responsibility for patient care	723	4.59 (0.66)	2.80		0.56	
14. Is reasonably accessible to patients	706	4.36 (0.77)	5.10		0.75	
15. Maintains confidentiality of patients	705	4.72 (0.52)	5.20	0.63		
16. Is accessible for appropriate communication to patients	724	4.48 (0.74)	2.70		0.75	
17. Communicates effectively with families	624	4.45 (0.74)	16.10	0.73		
18. Accepts responsibility for professional actions	702	4.60 (0.65)	5.60		0.60	
19. Responds appropriately in emergency situations	619	4.53 (0.70)	16.80		0.63	
20. Participates effectively as a member of the health care team	722	4.52 (0.76)	3.00		0.65	
21. Facilitates the learning of coworkers	684	4.42 (0.77)	8.10	0.61		
22. This doctor presents him or herself in a professional manner	741	4.64 (0.65)	0.40	0.75		
Cronbach's α (overall = 0.96) ^c				0.95	0.93	
Variance for each factor, % (total = 68.83)				59.22	5.14	4.47
UA = unable to access						
^a >20% indicated UA						
^b Extraction method: principal component analysis. - Rotation method: Varimax with Kaiser normalization rotation converged in 6 iterations.						
^c Generalizability coefficient, $E_p^2 = 0.82$						

Table 3 Descriptive statistics and item analysis for psychiatrists' medical colleagues and self-assessment

Research item	Medical colleagues, $\alpha = 0.98^a$			Self, $\alpha = 0.97^a$		
	<i>n</i>	Mean (SD)	UA, %	<i>n</i>	Mean (SD)	UA, %
1. Communicates effectively with patients	743	4.47 (0.63)	2.7	101	4.28 (0.59)	0
2. Is able to verbally communicate effectively with others	676	4.45 (0.65)	11.5	100	4.00 (0.64)	1.0
3. Communicates effectively with other health care professionals	753	4.44 (0.69)	1.4	101	3.99 (0.61)	0
4. Writes legibly	698	4.40 (0.65)	8.6	101	4.16 (0.60)	0
5. Within the range of services provided by this physician, he or she performs technical procedures skilfully	426	4.46 (0.61)	44.2	53	3.83 (0.67)	47.5
6. Selects diagnostic tests appropriately	626	4.31 (0.69)	18.1	98	3.67 (0.70)	3.0
7. Critically assesses diagnostic information	711	4.41 (0.64)	6.9	101	3.94 (0.58)	0
8. Makes the correct diagnosis following consultation	748	4.48 (0.62)	2.1	101	4.00 (0.62)	0
9. Selects appropriate treatments	754	4.46 (0.62)	1.3	101	4.09 (0.63)	0
10. Maintains quality records	589	4.30 (0.72)	22.9	101	3.70 (0.69)	0
11. Handles transfer of care appropriately	699	4.36 (0.68)	8.5	101	3.70 (0.69)	0
12. Provides a clear understanding about who is responsible for the continuing care of patients	727	4.41 (0.67)	4.8	101	3.84 (0.72)	0
13. Recognizes psychosocial aspects of illness	758	4.63 (0.59)	0.8	101	4.37 (0.66)	0
14. Maintains confidentiality of patients and their families	713	4.59 (0.57)	6.7	101	4.12 (0.71)	0
15. Coordinates care effectively for patients with other health care professionals and physicians	738	4.46 (0.66)	3.4	99	3.85 (0.65)	2.0
16. Manages patients with complex problems	745	4.42 (0.67)	2.5	101	4.04 (0.63)	0
17. Respects the rights of patients	748	4.52 (0.61)	2.1	101	4.29 (0.71)	0
18. Shows compassion for patients and their families	747	4.49 (0.66)	2.2	101	4.27 (0.68)	0
19. Collaborates with physician colleagues	751	4.44 (0.66)	1.7	100	3.83 (0.73)	1.0
20. Is involved with professional development	623	4.34 (0.71)	18.5	101	3.70 (0.82)	0
21. Accepts responsibility for own professional actions	715	4.47 (0.63)	6.4	101	4.08 (0.63)	0
22. Manages health care resources efficiently	633	4.30 (0.68)	17.1	101	3.62 (0.73)	0
23. Makes appropriate use of community resources for psychosocial aspects of care	706	4.48 (0.63)	7.6	101	3.79 (0.70)	0
24. Gives priority to urgent requests	699	4.41 (0.68)	8.5	99	3.86 (0.67)	2.0
25. Handles emergency situations effectively	620	4.42 (0.67)	18.8	99	3.85 (0.69)	2.0
26. Manages own stress effectively	604	4.23 (0.73)	20.9	100	3.68 (0.86)	1.0
27. Participates in a system of call to provide care for his or her own patients when unavailable	614	4.30 (0.73)	19.6	91	3.85 (0.73)	9.9
28. Recognizes his or her limitations	697	4.29 (0.70)	8.8	100	3.77 (0.75)	1.0
29. Handles requests for consultation in a timely manner	709	4.32 (0.73)	7.2	100	3.60 (0.68)	1.0
30. Advises referring physician if referral request is outside the scope of his or her practice	588	4.36 (0.66)	23.0	97	3.85 (0.77)	4.0
31. Assumes appropriate responsibility for patients	752	4.46 (0.64)	1.6	101	3.99 (0.69)	0
32. Provides timely information to referring physicians about mutual patients	677	4.38 (0.69)	11.4	101	3.60 (0.71)	0
33. Critically evaluates the medical literature to optimize clinical decision making	604	4.34 (0.68)	20.9	101	3.60 (0.74)	0
34. Facilitates the learning of medical colleagues and coworkers	683	4.34 (0.74)	10.6	95	3.75 (0.85)	5.9
35. Contributes to quality improvement programs and practice guidelines	527	4.20 (0.79)	31.0	89	3.25 (0.90)	11.9
36. Participates effectively as a member of the health care team	733	4.42 (0.69)	4.1	98	3.99 (0.73)	3.0
37. Exhibits professional and ethical behaviour towards physician colleagues	759	4.55 (0.64)	0.7	100	4.08 (0.73)	1.0
38. If a member of my own family needed care, I would rate this physician	758	4.49 (0.70)	0.8	Not in self-assessment		

^aGeneralizability coefficient, $Ep^2 = 0.81$; UA = unable to access

together in fairly coherent ways. As such, the factors provide the general direction for physician improvement, while the items themselves provide more specific guidance. Each physician received descriptive data (means and SDs) on the scales and individual items, for him or herself, as well as the group as a whole.

Last, we have evidence that the instruments are reliable both at an instrument and at an individual practitioner level. The internal consistency reliability analysis (Cronbach's alpha) suggests both the instrument and the scales are internally consistent. Further, the generalizability coefficient data indicate that the data provided to each physician were also stable across raters and comparable with that found in other studies.^{6–10,15} These data suggest that the mix of items and raters on the surveys is appropriate.

There are limitations in the study. Data testing was limited to psychiatrists in one Canadian province whose regulatory authority had mandated participation in the program. We do not know whether physicians in other parts of Canada or volunteer physicians would have similar performance profiles. Future research might well focus on replicating the present study in other jurisdictions (for example, other provinces). Further validity study could also be undertaken by studying the relation of survey scores with direct observation of clinical performance, for example.

This study, similar to all but one MSF study,¹⁵ permitted participating physicians to identify respondents. While this may introduce bias, the nature of medicine makes it difficult for anyone other than the physician to identify those medical colleagues and coworkers who can answer the questions. Further, in a province-wide program involving every physician every 5 years, other approaches to recruitment do not appear feasible at this time. This work is dependent on the professionalism of those involved. In the case of patients, however, we know from the generalizability coefficient analysis that 25 were required to reach an $E_p^2 = 0.78$ for 40 items, whereas 8 coworkers and medical colleagues produced $E_p^2 = 0.82$ and $E_p^2 = 0.81$, respectively, with 22 and 38 items, respectively. This suggests that patients' perspectives were more heterogeneous. However, there remains an issue about the appropriateness of physicians selecting their own patients in the MSF survey—future research can focus on a comparison of assessments by patients selected by the assessed physician and by someone else (for example, the researchers).

MSF is relatively new. With the expectation that physicians will achieve competence across numerous domains, MSF has a certain appeal as a way to inform physicians about professionalism, collegiality, and communication so that the professional can focus on selected aspects of practice behaviour. MSF appears to be a relatively inexpensive way of assessing

these competencies and assessing the changes physicians make based on the feedback received. While this study did not examine the use physicians made of their MSF data, this would be a legitimate scholarly inquiry. For example, each physician received personalized data as well as aggregate data. A follow-up study to determine how the physicians used their data, the changes they made as a result of the feedback, and their perceptions of this type of assessment is certainly warranted and was undertaken in other MSF work.^{6,11} For example, in a study employing 255 family physicians, we found that 66% initiated change to at least one aspect of their practice (such as, communication with patients and support of patients) as a consequence of individual feedback from MSF data.¹¹ At an institutional or provincial level, aggregate data for this group of physicians could be used as part of a needs assessment to guide educational program or institutional policy or procedures.

Conclusions

We believe our MSF instruments for psychiatrists provide a feasible way of assessing psychiatrists and of providing guided feedback on numerous competencies and behaviours. Our analysis provides evidence for the validity and reliability of the instruments. The questionnaires were developed with a regulatory authority as a quality improvement program. While the items focus on the needs of a regulatory authority, their breadth and scope may provide a base set of items on which to assess aspects of CanMEDS¹ competencies for other practising physicians and residents.

Funding and Support

The authors received funding from the College of Physicians and Surgeons of Alberta.

Acknowledgements

The authors thank the physicians who served on the steering committee, Dr Martin Atkinson, Dr Decker Butzner, Dr William Dickert, Dr Alan Gordon, Dr Alexander Leung, Dr Raymond Lewkonja, Dr Ross McLeod, and Dr Neil Roberts. We thank John Swiniarski and Bryan Ward, College of Physicians and Surgeons of Alberta.

References

1. Frank J, editor. The CanMEDS 2005 physician competence framework [Internet]. Ottawa (ON): Royal College of Physicians and Surgeons of Canada; 2005 [cited 2007 Jun 18]. Available from: http://rcpsc.medical.org/canmeds/CanMEDS2005/CanMEDS2005_e.pdf.
2. Accreditation Council for Graduate Medical Education. ACGME Outcome Project [Internet]. Accreditation Council for Graduate Medical Education Chicago (IL): 2006 [cited 2007 Jun 18]. Available from: <http://www.acgme.org/Outcome/>.
3. General Medical Council. Good medical practice 2006 [Internet]. London (GB): General Medical Council; 2006 [cited 2007 Jun 18]. Available from: http://www.gmc-uk.org/guidance/good_medical_practice/GMC_GMP.pdf.
4. Swick S, Hall S, Beresin E. Assessing the ACGME competencies in psychiatry training programs. *Acad Psychiatry*. 2006;30:330–351.
5. Bandiera G, Sherbino J, Frank J, editors. The CanMEDS assessment tools handbook: an introductory guide for assessment methods for the CanMEDS

- competencies. Ottawa (ON): Royal College of Physicians and Surgeons of Canada; 2006.
6. Archer J, Norcini J, Southgate L, et al. Mini-PAT (peer assessment tool): a valid component of a national assessment programme in the UK? *Adv Health Sci Educ Theory Pract.* 2006;Oct 12. DOI: 10.1007/s10459-006-9033-3.
 7. Hall W, Violato C, Lewkonja R, et al. Assessment of physician performance in Alberta: the Physician Achievement Review. *CMAJ.* 1999;61:52-57.
 8. Lipner RS, Blank LL, Leas BF, et al. The value of patient and peer ratings in recertification. *Acad Med.* 2002;77(S):64-66.
 9. Lockyer JM, Violato C, Fidler H. A multisource feedback program for anesthesiologists. *Can J Anesth.* 2006;53(1):33-39.
 10. Violato C, Lockyer JM, Fidler H. Multi source feedback: a method of assessing surgical practice. *BMJ.* 2003;326:546-548.
 11. Fidler H, Lockyer J, Toews J, et al. Changing physicians' practices: the effect of individual feedback. *Acad Med.* 1999;74:702-714.
 12. College of Physicians and Surgeons of Alberta. Physician Achievement Review Program [Internet]. Edmonton (AB): College of Physicians and Surgeons of Alberta; 2005 [cited 2007 Jun 18]. Available from: <http://www.par-program.org/PAR-Inst.htm>.
 13. Lockyer J, Violato C. An examination of the appropriateness of using a common peer assessment instrument to assess physician skills across specialties. *Acad Med.* 2004;79(10 Suppl):S5-S8.
 14. Violato C, Lockyer J. Self and peer assessment of pediatricians, psychiatrists and medicine specialists: implications for self-directed learning. *Advances in Health Science Education.* 2006;11(3):235-244.
 15. Ramsey PG, Wenrich MD, Carline JD, et al. Use of peer ratings to evaluate physician performance. *JAMA.* 1993;269(13):1655-1660.

Manuscript received July 2007, revised, and accepted January 2008.

¹Professor and Director of the Medical Education and Research Unit, Faculty of Medicine, University of Calgary, Calgary, Alberta.

²Associate Dean of Continuing Medical Education and Professional Development, Associate Professor, Department of Community Health Sciences, Faculty of Medicine, University of Calgary, Calgary, Alberta.

³Research Associate, Office of Continuing Medical Education and Professional Development, Faculty of Medicine, University of Calgary, Calgary, Alberta.

Address for correspondence: Dr C Violato, Department of Community Health Sciences, University of Calgary, 3330 Hospital Drive NW, Calgary, AB T2N 4N1; violato@ucalgary.ca

Résumé : L'évaluation des psychiatres dans la pratique par une rétroaction multi-sources

Objectif : Évaluer la faisabilité et les preuves de la fiabilité et de la validité d'un ensemble de questionnaires pour les psychiatres, étant donné que la rétroaction multi-sources (RMS) ou évaluation 360° permet aux collègues médicaux, aux collègues en général et aux patients de fournir une rétroaction sur les compétences, afin d'accroître l'amélioration du médecin dans des directions voulues.

Méthode : Des questionnaires de 40, 22, 38, et 37 items ont été élaborés pour évaluer les psychiatres par 25 patients, 8 collègues, 8 psychiatres collègues, et les psychiatres eux-mêmes, respectivement, à l'aide d'une échelle d'accord pourvue d'une catégorie « incapable d'évaluer ». Les items concernaient les compétences liées aux aptitudes à la communication, au professionnalisme, à la collégialité, et à l'autogestion. La faisabilité était évaluée par les taux de réponse à chaque instrument. La validité était évaluée par une table de spécifications, le pourcentage de participants incapables d'évaluer le psychiatre pour chaque item, et des analyses factorielles exploratoires pour déterminer quels items se regroupaient dans les échelles. La fiabilité était évaluée par le coefficient alpha de Cronbach et le coefficient de généralisabilité.

Résultats : Un échantillon de 101 psychiatres a fourni des données. Un total de 2 456 patients (24,32/25 par psychiatre), 744 collègues psychiatres (7,37/8 par psychiatre), 764 collègues (7,56/8 par psychiatre), et 101 formulaires autodéclarés ont été analysés. La fiabilité de la cohésion interne générale des instruments était un alpha de Cronbach de 0,98, 0,96, et 0,98 pour les questionnaires des patients, des collègues et des collègues médicaux, respectivement. Le coefficient de généralisabilité pour les patients, collègues et collègues médicaux était de 0,78, 0,82, et 0,81, respectivement.

Conclusion : Il est possible d'élaborer un programme de RMS faisable pour les psychiatres avec fiabilité et validité éprouvées, qui peut procurer une rétroaction sur les compétences cliniques fondamentales.