

## Effectiveness of a quality-improvement program in improving management of primary care practices

Joachim Szecsenyi MD MSc, Stephen Campbell PhD, Bjoern Broge MBA, Gunter Laux PhD, Sara Willms MSc, Michel Wensing PhD, Katja Goetz PhD

### Competing interests:

Bjoern Broge and Sara Willms are employed by the AQUA-Institute, which disseminates the European Practice Assessment in Germany. Joachim Szecsenyi is its director and a stockholder. No competing interests declared by Stephen Campbell, Gunter Laux, Michel Wensing or Katja Goetz.

This article has been peer reviewed.

### Correspondence to:

Dr. Katja Goetz; katja.goetz@med.uni-heidelberg.de

CMAJ 2011. DOI:10.1503/cmaj.110412

### ABSTRACT

**Background:** The European Practice Assessment program provides feedback and outreach visits to primary care practices to facilitate quality improvement in five domains (infrastructure, people, information, finance, and quality and safety). We examined the effectiveness of this program in improving management in primary care practices in Germany, with a focus on the domain of quality and safety.

**Methods:** In a before–after study, 102 primary care practices completed a practice assessment using the European Practice Assessment instrument at baseline and three years later (intervention group). A comparative group of 102 practices was included that completed their first assessment using this instrument at the time of the intervention group's second assessment. Mean scores were based on the proportion of indicators for which a positive response was achieved by all of the practices, on a scale of 0 to 100.

**Results:** We found significant improvements in all domains between the first and second

assessments in the intervention group. In the domain of quality and safety, improvements in scores (mean scores were based on the proportion of indicators for which a positive response was achieved by all of the practices, on a scale of 0 to 100) were observed in the following dimensions: complaint management (from a mean score of 51.2 at first assessment to 80.7 at second assessment); analysis of critical incidents (from 79.1 to 89.6); and quality development, quality policy (from 40.7 to 55.6). Overall scores at the time of the second assessment were significantly higher in the intervention group than in the comparative group.

**Interpretation:** Primary care practices that completed the European Practice Assessment instrument twice over a three-year period showed improvements in practice management. Our findings show the value of the quality-improvement cycle in the context of practice assessment and the use of established organizational standards for practice management with the European Practice Assessment.

A variety of quality-improvement initiatives in health care management have been implemented in most health care systems.<sup>1,2</sup> Countries such as Australia, New Zealand, the United Kingdom and the United States have a long tradition in, and established standards for, quality management in primary care. In Canada, such initiatives in primary care are in their infancy, despite support by the federal Primary Health Care Transition Fund since 2000.<sup>3</sup> In Ontario, a comprehensive book was recently issued that provides recommendations on practice management and clinical indicators for improving quality in primary care settings in the province.<sup>4</sup> The indicators were adopted and refined from the Royal New Zealand College of General Practitioners, the Royal Australian College of General Practitioners' National Expert

Committee on Standards for General Practices, the TOPAS–Europe Association European Practice Assessment, the United Kingdom's Quality and Outcomes Framework and the Canadian Institute for Health Information.

In Germany, similar developments took place. In 2005, the German government stipulated that health care providers implement a system of annual assessment of quality management.<sup>5</sup> One of the systems available to practices is the European Practice Assessment ([www.epa-qm.de](http://www.epa-qm.de)), a validated instrument based on quality indicators for assessing practice management.<sup>6</sup> The five key domains of the European Practice Assessment instrument and their respective dimensions are described in Figure 1.

The European Practice Assessment is used to help general practices both assess and improve

their quality of practice management set against predefined criteria with embedded quality indicators. The improvement process is ongoing (e.g., through plan–do–study–act cycles<sup>7</sup>): each step is reviewed and redesigned with a view to improving the quality of the end product, thereby fostering continuous improvement.<sup>8</sup> The multifaceted strategy has three essential components: assessment and feedback using validated instruments based on quality indicators; external support through an outreach educational visit by a trained visitor (physician or nurse) to support the practice in improving areas of management identified by the practice;<sup>9</sup> and formal accreditation by an external organization.

Three requirements have to be fulfilled by practices to receive accreditation: achieve a positive response for more than 50% of the indicators; meet predefined safety indicators (e.g., the vaccination status of staff regarding hepatitis B vaccination is recorded and medical equipment is checked regularly according to national regulations); and highlight areas for continuous quality improvement.

Accreditation is one method for assessing and benchmarking the performance of general practice care across a broad range of clinical and organizational domains. It describes a formal process of self-assessment and external and independent peer review to encourage best management practice and can result in recommendations for

continuous improvement of safety and quality in the practice.<sup>8</sup> Practice accreditation can be used for different purposes: quality control, regulation, quality improvement, providing data on performance, and marketing.<sup>8</sup> In Germany, it is used for quality improvement, leading to a certificate.

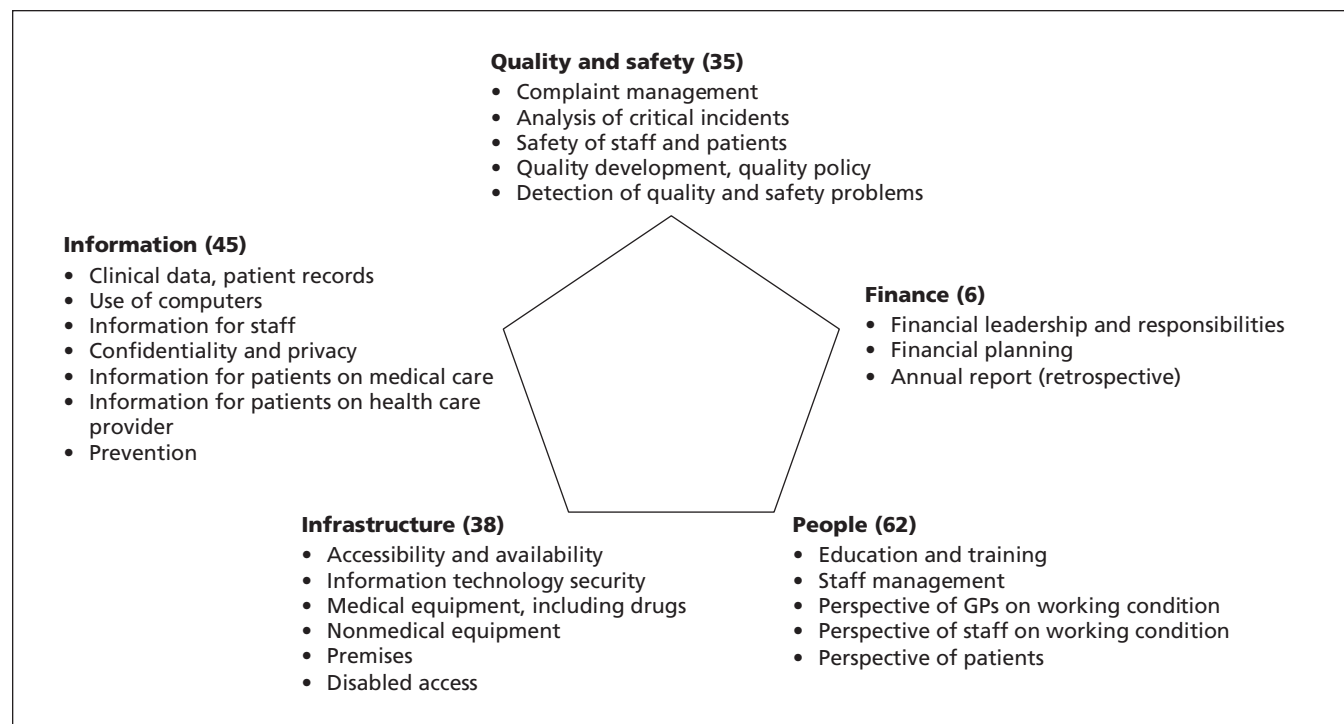
We conducted a study to determine whether improvements in practice management occurred in general practices that completed the European Practice Assessment twice over three years, compared with general practices that completed the European Practice Assessment once. We focused our analysis on the domain of quality and safety, expecting an association between practice organization and quality improvement.<sup>10,11</sup> We hypothesized that the initial use of the European Practice Assessment and reassessment with it three years later would result in improved scores in the dimensions and indicators within the domain of quality and safety.

## Methods

### Study practices

We included general practices in Germany that had completed the European Practice Assessment as part of a quality-improvement program for ambulatory care between May 2004 and January 2009.

For the intervention group, we selected 110 practices that underwent their first assessment



**Figure 1: The domains and dimensions (and number of indicators) of the European Practice Assessment instrument used to measure the quality of practice management in primary care practices. For an example of how the pentagon shape is used to provide feedback to individual practices, see Appendix 1 (available at [www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.110412/-/DC1](http://www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.110412/-/DC1)). GP = general practitioner.**

between May 2004 and November 2007 and whose staff completed the European Practice Assessment instruments. These practices underwent a second assessment between June 2007 and January 2009. The interval between assessments for each participating practice was 36 months. We excluded eight practices that had different intervals, which left 102 practices in the intervention group.

During the period when the intervention practices were undergoing their second assessment, 209 general practices underwent their first assessment with the European Practice Assessment program but had not yet had a second assessment. These practices were considered for inclusion in the comparative group.

Because both groups differed on several measures, we used a matched pair design based on propensity scores.<sup>12,13</sup> The weights for the propensity score were calculated on the basis of binary logistic regression for the following covariates: practice type (solo v. group), location (urban v. rural), number of general practitioners per practice, number of nurses per practice and mean number of patients per practice seen in a quarter. We used the “Matching” package of the R Project (version 2.11.1) to perform the matching.<sup>14,15</sup> We used the “MatchBalance” function of the package to evaluate the matching results, because this function not only examines differences in proportions or means, it also accounts

for potential differences across the entire distribution.<sup>14</sup> After matching, we had 102 comparable practices in the comparative group (Figure 2).

### Practice assessment

The application of the European Practice Assessment is coordinated by the AQUA-Institute and is described in detail elsewhere.<sup>6</sup> In brief, the self-administered questionnaire was sent to all practice staff (including the practice manager where employed) and individual general practitioners. Once completed questionnaires were returned, a practice visit was conducted by a specially trained visitor. The visitor first assessed the practice against a standardized checklist and then conducted a structured interview with the general practitioner who had management responsibilities or the practice manager. The visit finished with a feedback session on the preliminary outcomes of the assessment. The data file Visotool was used to show the results of the assessment and provide an anonymous comparison between the practice's scores and all other practices that had undertaken the assessment, as a catalyst for quality improvement. (An example of a Visotool figure is presented in Appendix 1, available at [www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.110412/-/DC1](http://www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.110412/-/DC1)).

### Statistical analysis

We used *z* scores to compare the mean scores for all domains of the European Practice Assessment instrument, and specifically for the dimensions and indicators in the “quality and safety” domain, in the intervention and comparative groups. (Mean scores were based on the proportion of indicators for which a positive response was achieved by all of the practices, on a scale of 0 to 100.) Establishing a *z* score is a procedure for normalizing aggregated data and is often used when comparing practices.<sup>16</sup> We calculated 95% confidence intervals for the differences in scores in the intervention group (between the first and second assessments) and for the between-group differences in scores (between the second assessment in the intervention group and the first assessment in the comparative group). We used the  $\chi^2$  test to compare characteristics of the practices in the intervention and comparative groups. An  $\alpha$  level of  $p \leq 0.05$  was used for tests of statistical significance.

### Ethics approval

Ethical approval was not required because we used data available from the routine implementation of a quality-improvement program in ambulatory care in Germany.<sup>5</sup> All European Practice Assessment questionnaires and the information from the trained visitors were anonymized for

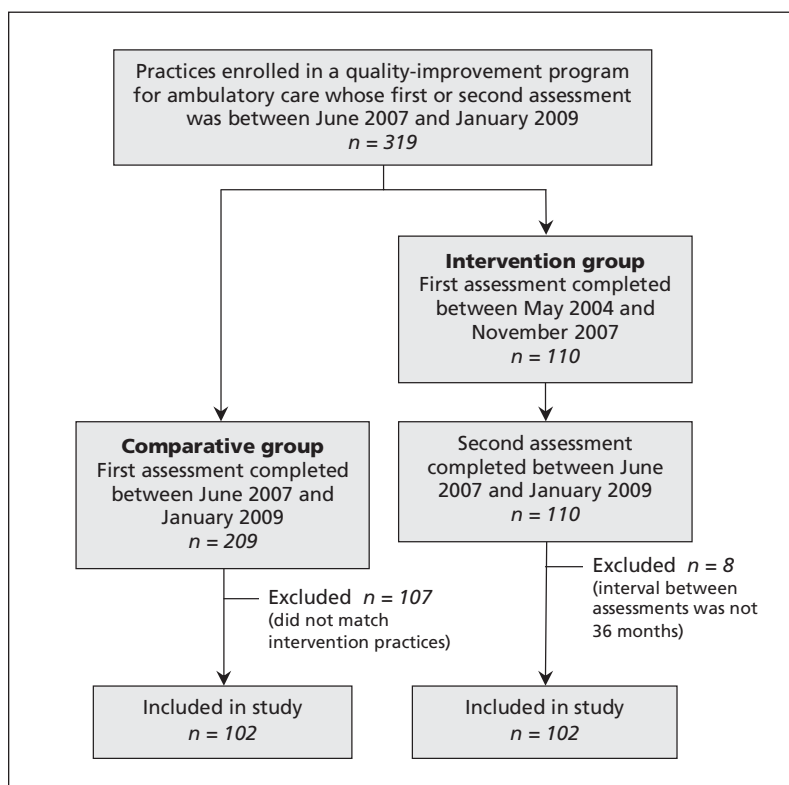


Figure 2: Selection of primary care practices for the intervention and comparative groups.

data analysis in our study. No additional information or data from patients or staff were collected.

## Results

The practices included in the intervention and comparative groups did not differ significantly in the relevant covariates of practice characteristics (Table 1).

The descriptive results for European Practice Assessment are described elsewhere.<sup>17</sup> We observed an improvement in quality management in the intervention group, with a change in score of 7.0 in the five domains overall. The change in score was 10.1 in the domain of quality and safety, which is analyzed in detail in this study.

## Quality and safety

Changes in the scores for the dimensions in the domain of quality and safety are shown in Table 2. A significant improvement occurred in the intervention group in the dimensions “complaint management,” “analysis of critical incidents” and “quality development, quality policy.” Compared with the comparative group at first assessment, the intervention group at reassessment had better overall scores for the dimensions “detection of quality and safety problems,” “complaint management,” “analysis of critical incidents” and “quality development, quality policy.”

Changes in scores for each of the indicators in the domain of quality and safety are presented in Table 3. Significant improvements were observed in the intervention group, especially in the dimensions “complaint management” (e.g., “there is a patient complaint procedure”), “analysis of critical incidents” (e.g., “there is a critical incident register”) and “qual-

ity development, quality policy” (e.g., “patient satisfaction is monitored regularly”).

## Interpretation

In our study, primary care practices that had undergone the European Practice Assessment process showed overall improvements in practice management at the time of the second assessment. In the domain of quality and safety, we

**Table 1:** Characteristics of primary care practices included in the study\*

Characteristic	Intervention group n = 102	Comparative group n = 102
Type, no. (%)		
Solo	50 (49.0)	54 (52.9)
Group	51 (50.0)	48 (47.1)
Data missing	1 (1.0)	–
Location, no. (%)		
Urban	41 (40.2)	39 (38.2)
Rural	59 (57.8)	61 (59.8)
Data missing	2 (2.0)	2 (2.0)
No. of general practitioners per practice		
Mean (SD)	1.7 (0.8)	1.6 (0.8)
Range	1–4	1–5
No. of nurses per practice		
Mean (SD)	5.8 (2.6)	5.6 (2.1)
Range	1–17	1–13
No. of patients per practice seen in a quarter, mean (SD)	1774.8 (745.9)	1703.9 (758.9)
Note: SD = standard deviation. *Characteristics are shown for the period 2007–2009, during which the intervention practices underwent their second assessment and the comparative practices underwent their first assessment.		

**Table 2:** Mean scores\* for the dimensions in the quality and safety domain of the European Practice Assessment instrument

Dimension	No. of indicators	Intervention group			Score for comparative group†	Between-group difference in scores‡ (95% CI)
		Score at first assessment	Score at second assessment	Change in score (95% CI)		
Complaint management	6	51.2	80.7	29.5 (20.7 to 38.4)	66.5	14.2 (7.4 to 21.0)
Analysis of critical incidents	5	79.1	89.6	10.5 (4.5 to 16.4)	83.9	5.7 (1.2 to 10.2)
Safety of staff and patients, hygiene, infection control	12	85.8	89.1	3.3 (–0.3 to 6.3)	89.0	0.1 (–0.5 to 0.7)
Quality development, quality policy	7	40.7	55.6	14.9 (7.9 to 21.8)	40.8	14.8 (7.9 to 21.7)
Detection of quality and safety problems	5	86.4	89.6	3.2 (–0.2 to 6.6)	77.8	11.8 (5.5 to 18.1)
All	35	72.3	82.4	10.1 (4.2 to 15.9)	73.9	8.5 (3.1 to 13.9)
Note: CI = confidence interval. *Mean scores are on a scale of 0 to 100 and are based on the proportion of indicators for which a positive response was achieved by all of the practices. †Mean scores at first assessment. ‡Comparison between mean scores at second assessment for intervention practices and mean scores at first assessment for comparative practices.						

**Table 3:** Mean scores\* for the indicators in the quality and safety domain of the European Practice Assessment instrument (part 1 of 2)

Dimension; indicator	Intervention group			Score for comparative group†	Between-group difference in scores‡ (95% CI)	
	Score at first assessment	Score at second assessment	Change in score (95% CI)			
<b>Complaint management</b>						
Patient complaints are acted upon	69.6	91.7	22.1 (14.0 to 30.1)	84.3	7.4	(2.3 to 12.5)
There is a patient complaint procedure	43.0	88.0	45.0 (35.3 to 54.6)	63.4	24.6	(16.2 to 33.0)
There is a suggestion box for patients	7.8	43.1	35.3 (26.0 to 44.6)	20.6	22.5	(14.4 to 30.6)
All patient complaints are analyzed and discussed	84.7	100.0	15.2 (8.0 to 22.2)	93.8	6.2	(1.5 to 10.9)
An example of a patient complaint that has been discussed is available	46.1	95.1	49.0 (39.3 to 58.7)	84.3	10.8	(4.8 to 16.8)
All patient complaints are registered	55.9	66.7	10.7 (4.7 to 16.7)	52.6	14.1	(7.3 to 20.8)
<b>Analysis of critical incidents</b>						
There is a register for critical incidents	30.7	58.8	28.1 (19.4 to 36.8)	46.1	12.7	(6.2 to 19.2)
Examples of analyses of critical incidents are available	94.1	97.1	2.9 (-0.3 to 6.2)	94.1	3.0	(-0.3 to 6.3)
Critical incidents are analyzed and discussed	86.3	96.1	9.8 (4.0 to 15.6)	96.0	0.1	(-0.5 to 0.7)
Critical incidents are acted upon	93.1	98.0	4.9 (0.6 to 9.0)	94.1	3.9	(0.1 to 7.6)
Examples of action taken because of a critical incident are available	91.2	98.0	6.8 (1.9 to 11.8)	89.2	8.8	(3.3 to 14.3)
<b>Safety of staff and patients, hygiene, infection control</b>						
The practice has a container for used equipment in the consultation/examination rooms	83.1	87.2	4.1 (0.2 to 7.9)	92.6	-5.4	(-9.8 to -1.0)
There is a written infection control protocol for the prevention of contamination of staff	97.1	99.0	1.9 (-0.7 to 4.6)	98.0	1.0	(-0.9 to 2.9)
The practice has a leak-proof container in the consultation/examination rooms	83.1	98.8	15.7 (8.6 to 22.7)	92.8	6.0	(1.4 to 10.6)
The practice has hygienic hand wash and disinfection facilities and drying facilities in the consultation/examination rooms	73.0	74.9	1.8 (-0.7 to 4.4)	79.9	-5.0	(-9.2 to -0.07)
The practice has a container for sharps in the consultation/examination rooms	90.5	98.0	7.5 (2.4 to 12.6)	98.2	-0.2	(-10.7 to -0.06)
The hygiene protocol covers disinfection of clinical equipment	91.8	97.6	5.8 (1.3 to 10.4)	98.6	-1.0	(-2.9 to -0.09)
The hygiene protocol covers the disposal of sharps and contaminated material	83.6	91.7	8.1 (2.7 to 13.3)	85.5	6.2	(1.5 to 10.8)
The hygiene protocol covers the use of protective equipment	80.3	87.8	7.5 (2.3 to 12.6)	91.4	-3.6	(-7.2 to 0.0)
The hygiene protocol covers the disposal of used instruments	95.1	96.2	1.1 (-0.9 to 3.1)	97.1	-0.9	(-2.7 to -0.09)
The hygiene protocol covers when and how to use sterile instruments	89.8	78.5	-11.3 (-17.5 to -5.2)	83.8	-5.3	(-9.6 to -0.09)
The vaccination status of staff regarding hepatitis B vaccination is recorded	97.5	94.1	-3.3 (-6.8 to 0.1)	86.3	7.8	(2.5 to 13.0)
There is a written hygiene protocol	75.9	81.4	5.4 (1.0 to 9.8)	64.0	17.4	(10.0 to 24.7)
<b>Quality development, quality policy</b>						
Patient satisfaction is monitored regularly	35.3	82.8	47.4 (37.7 to 57.1)	25.6	57.2	(47.6 to 66.8)
Practice staff have protected time to discuss the future strategy of the practice in the past 12 months	51.7	47.1	-4.6 (-8.7 to -0.5)	40.2	6.2	(1.5 to 10.9)

**Table 3:** Mean scores\* for the indicators in the quality and safety domain of the European Practice Assessment instrument (part 2 of 2)

Dimension; indicator	Intervention group			Score for comparative group†	Between-group difference in scores‡ (95% CI)
	Score at first assessment	Score at second assessment	Change in score (95% CI)		
There is a patient forum or a patient participation group	12.8	2.0	-10.8 (-16.8 to -4.7)	2.9	-0.9 (-14.5 to -3.4)
Quality-improvement targets have been set in the last year	80.0	92.8	12.8 (6.3 to 19.3)	87.6	5.2 (0.08 to 9.5)
At least one clinical audit has been conducted in the last year	42.3	71.0	28.7 (20.0 to 37.4)	50.3	20.7 (12.8 to 28.6)
The annual report includes a quality report	6.9	16.7	9.7 (4.0 to 15.5)	7.8	8.9 (3.3 to 14.4)
A team meeting is held at least monthly and a written record is made available to all staff	55.8	77.3	21.5 (13.5 to 29.4)	71.5	5.8 (1.3 to 10.3)
<b>Detection of quality and safety problems</b>					
Sphygmomanometers are calibrated regularly according to national regulations	86.3	92.2	5.9 (1.3 to 10.4)	86.3	5.9 (1.3 to 10.5)
Electronic medical equipment (e.g., defibrillator) is checked regularly according to national regulations	94.1	87.3	-6.8 (-11.7 to -1.9)	78.4	8.9 (3.4 to 14.4)
Medical equipment (e.g., ophthalmoscopes, weigh scales, instruments) is checked regularly according to national regulations	89.2	92.2	2.9 (-0.3 to 6.2)	79.4	12.8 (6.3 to 19.3)
The safety equipment is checked regularly according to national regulations	92.2	94.1	2.0 (-0.7 to 4.6)	86.3	7.8 (6.3 to 19.3)
National ergonomic standards are adhered to	70.3	82.3	12.0 (5.7 to 18.4)	58.4	23.9 (15.6 to 32.2)

Note: CI = confidence interval.  
 \*Mean scores are on a scale of 0 to 100 and are based on the proportion of indicators for which a positive response was achieved by all of the practices.  
 †Mean scores at first assessment.  
 ‡Comparison between mean scores at second assessment for intervention practices and mean scores at first assessment for comparative practices.

observed significant improvements in three of the five dimensions (“complaint management,” “analysis of critical incidents” and “quality development, quality policy”).

Two aspects are important when interpreting our results. An increase of 10.1 in the overall score for the domain of quality and safety could indicate well-organized and safe practices, which is important from the perspective of patients. From the perspective of practices, our study showed larger changes than studies did using clinical indicators.<sup>18</sup>

Similar to another evaluation of a quality-improvement program,<sup>19</sup> we used repeated measurement to evaluate the effect of the assessment process. However, we included a comparative group. Moreover, the European Practice Assessment involves both quality assessment and quality improvement, for example in relation to the analysis of critical incidents, complaint management, and the detection of problems in quality and safety. This shows that patient safety is also an integral part of the European Practice Assessment, as one key component of quality of care.<sup>20</sup>

Quality improvement requires a combination of educational, organizational and financial approaches, using both intrinsic motivation and external incentives.<sup>21</sup> For example, an Australia quality-improvement initiative “the Bettering the Evaluation and Care of Health program” showed that certification was positively associated with consultation behaviour and patient management.<sup>22</sup>

Quality of care is a broad concept, which requires a mix of objective and subjective measures.<sup>23</sup> One method is formative accreditation, which combines a broad set of quality measures and various improvement activities such as audit and feedback,<sup>10</sup> and educational outreach visits.<sup>11</sup> These strategies have been shown to be effective, and the context of practice accreditation might further add to their impact. Many methods for quality improvement have been shown to have only short-term effects at best. For instance, pay-for-performance has been introduced in many countries and has shown short-term improvements, but the evidence has not been compelling and most incentives were attached to meeting a

set target rather than quality improvement.<sup>24</sup> Audit and feedback have been shown to have similar short-term effects.<sup>10</sup> The same applies to feedback given to primary care providers on patients' evaluations of care.<sup>25</sup> This may suggest that such approaches need to be used as one part of a multiple-component strategy for quality improvement rather than used in isolation.<sup>26</sup> Our findings support the use and role of formative accreditation as one part of such a strategy.

### Limitations

Our study has limitations. First, the sample of participating practices was small and may not have been representative of general practices in Germany because we involved only practices that elected to use the European Practice Assessment to fulfil their obligation of following the national directive for quality management in office-based practices, which introduced a selection bias. Second, the improvement observed at the second assessment may have been attributed in part to the Hawthorne effect. Third, we used propensity scores to match comparable practices in the intervention and comparative groups, but not for statistical adjustment. Lastly, because this was an exploratory study, changes were only descriptive in nature and will need to be confirmed in further targeted studies. A well-organized, multifaceted quality management program motivates practices to change.<sup>27</sup> However, until there is evidence of the impact on patient outcomes, it will be unclear how to weigh the benefits and risks of interventions such as the European Practice Assessment.

### Conclusion

Primary care practices that completed the European Practice Assessment twice over a three-year period showed improvements in practice management. Our findings show the value of the quality-improvement cycle in the context of practice assessment. Unlike other assessments against pre-defined standards such as the Care Quality Commission or the Quality Outcome Framework in the United Kingdom<sup>18</sup> or quality initiatives in Ontario,<sup>4</sup> which all have target standards, quality improvement is an integral part of the European Practice Assessment. Our findings suggest that there is an intrinsic benefit to practices undertaking formative accreditation in terms of practice-specific quality improvement. The benchmark assessment motivates practices to improve their performance. Our study provides a better understanding of how accreditation can help to improve quality of care by enabling practices to both meet set standards and to identify and address internal priorities for quality improvement.

### References

1. World Health Organization. *Quality and accreditation in health services. A global review*. Geneva (Switzerland): The Organization; 2003.
2. Greenfield D, Braithwaite J. Health sector accreditation research: a systematic review. *Int J Qual Health Care* 2008;20:172-83.
3. Hutchison B. Cracks in the foundation: the precarious state of Canada's primary care infrastructure. *Healthc Policy* 2007;3:10-6.
4. Levitt C, Hilt L. *Quality in family practice. Book of tools*. Hamilton (ON): McMaster Innovation Press; 2010.
5. Quality management directive on medical and dental care [article in German]. Berlin (Germany): Federal Joint Committee (G-BA); 2005. Available: [www.g-ba.de/downloads/62-492-3/RL\\_QM-Vertragsarzt-2005-10-18.pdf](http://www.g-ba.de/downloads/62-492-3/RL_QM-Vertragsarzt-2005-10-18.pdf) (accessed 2011 June 10).
6. Engels Y, Dautzenberg M, Campbell S, et al. Testing a European set of indicators for the evaluation of the management of primary care practices. *Fam Pract* 2006;23:137-47.
7. Deming EW. *Out of crisis*. Cambridge (MA): Massachusetts Institute of Technology, Centre for Advanced Engineering Study; 1986.
8. Buetow SA, Wellingham J. Accreditation of general practice: challenges and lessons. *Qual Saf Health Care* 2003;12:129-35.
9. van den Hombergh P, Grol R, van den Hoogen HJM, et al. Practice visits as a tool in quality improvement: acceptance and feasibility. *Qual Health Care* 1999;8:167-71.
10. Jamtvedt G, Young JM, Kristoffersen DT, et al. Audit and feedback: effects on professional practice and health care outcomes [review]. *Cochrane Database Syst Rev* 2006;(2):CD000259.
11. O'Brien MA, Rogers S, Jamtvedt G, et al. Educational outreach visits: effects on professional practice and health care outcomes [review]. *Cochrane Database Syst Rev* 2007;(4):CD000409.
12. Nam JM. Establishing equivalence of two treatments and sample size requirements in matched-pairs design. *Biometrics* 1997;53:1422-30.
13. Rosenbaum PR, Rubin DB. The central role of the Propensity Score in observational studies for causal effects. *Biometrika* 1983;70:41-55.
14. Sekhon JS. Multivariate and propensity score matching software with automated balance optimization: the Matching package for R. *J Stat Softw* 2011;42(7). Available: [www.jstatsoft.org/v42/i07/paper](http://www.jstatsoft.org/v42/i07/paper) (accessed 2011 Oct. 17).
15. *The R Project for Statistical Computing*. Wien (Austria): Institute for Statistical Mathematics; 2011. Available: [www.r-project.org](http://www.r-project.org) (accessed 2011 Mar. 9).
16. Goderis G, Borgermans L, Heyrman J, et al. Monitoring modifiable cardiovascular risk in type 2 diabetes care in general practice. The use of an aggregated z-score. *Med Care* 2010;48:589-95.
17. Szecsenyi J, Broge B, Willms S, et al. Benchmarking in ambulatory care practices — the European Practice Assessment (EPA) [article in German]. *Z Evid Forbild Qual Gesundh wesen* 2011;105:404-7.
18. Campbell S, Reeves D, Kontopantelis E, et al. Quality of primary care in England with the introduction of pay for performance. *N Engl J Med* 2007;357:181-90.
19. Mandel D, Amital H, Zimlichman E, et al. Quality assessment program in primary care clinics: a tool for quality improvement. *Int J Qual Health Care* 2004;16:175-80.
20. Baker GR, Norton PG, Flintoft V, et al. The Canadian adverse events study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678-86.
21. Rhydderch M, Edwards A, Elwyn G, et al. Organizational assessment in general practice: a systematic review and implications for quality improvement. *J Eval Clin Pract* 2005;11:366-78.
22. Miller G, Britt H, Pan Y, et al. Relationship between general practitioner certification and characteristics of care. *Med Care* 2004;42:770-8.
23. Greenhalgh T, Heath I. Measuring quality in the therapeutic relationship. Part 1: objective approaches. *Qual Saf Health Care* 2010;19:475-8.
24. Christianson J, Leatherman S, Sutherland K. *Financial incentives, healthcare providers and quality improvements: a review of the evidence*. London (UK): The Health Foundation; 2008.
25. Vingerhoets E, Wensing M, Grol R. Feedback of patients' evaluations of general practice care: a randomised trial. *Qual Health Care* 2001;10:224-8.
26. Ferlie EB, Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *Milbank Q* 2001;79:281-315.
27. Grimshaw J, McAuley LM, Bero LA, et al. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Qual Saf Health Care* 2003;12:298-303.

**Affiliations:** From the Department of General Practice and Health Services Research (Szecsenyi, Campbell, Laux, Wensing, Goetz), University of Heidelberg, Heidelberg, Germany; the Primary Care Research Group (Campbell), University of Manchester, Manchester, United Kingdom; the AQUA-Institute for Applied Quality Improvement and Research in Health Care (Szecsenyi, Broge, Willms), Goettingen, Germany; the Scientific Institute for Quality of Health Care, Radboud University Nijmegen Medical Centre (Wensing), Nijmegen, the Netherlands

**Contributors:** All of the authors contributed to the study concept. Joachim Szecsenyi, Michel Wensing and Katja Goetz designed the study. Bjoern Broge and Sara Willms collected the data. Joachim Szecsenyi and Katja Goetz conducted the initial data analyses and wrote the first draft of the manuscript. Stephen Campbell, Gunter Laux and Michel Wensing contributed to the interpretation of the data. All of

the authors reviewed and revised the manuscript and approved the final version submitted for publication.

**Funding:** This study was funded by the Department of General Practice and Health Services Research, University of Heidelberg.

**Acknowledgements:** The authors thank the primary care practices and outreach visitors who participated in the study. The authors also thank Marc Brodowski and Mohamed Lambarki from the European Practice Assessment management team at the AQUA-Institute for providing management data, the Stiftung Praxissiegel e.V., which serves as an independent certification or accreditation body in Germany, and the European Association for Quality in General Practice / Family Medicine (EQuiP) for supporting the European Practice Assessment project and the development of the indicators.