

Effect of e-Learning and Repeated Performance Feedback on Spirometry Test Quality in Family Practice: A Cluster Trial

Tjard R. Schermer, MSc, PhD¹

Reinier P. Akkermans, MSc¹

Alan J. Crockett, MD, PhD²

Marian van Montfort, MD³

Joke Grootens-Stekelenburg⁴

Jim W. Stout, MD, PhD⁴

Willem Pieters, MD, PhD⁵

¹Department of Primary and Community Care, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

²Discipline of General Practice, The University of Adelaide, Adelaide, Australia

³QUARTZ Integrated Care Support Service, Helmond, The Netherlands

⁴Department of Pediatrics, University of Washington, Seattle, Washington

⁵Department of Pulmonary Medicine, Elik-erliek Hospital, Helmond, The Netherlands



Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Tjard R. Schermer, MSc, PhD
Department of Primary and Community
Care (117-ELG)
Radboud University Nijmegen Medical
Centre
PO Box 9101
6500 HB Nijmegen
The Netherlands
t.schermer@elg.umcn.nl

ABSTRACT

PURPOSE Spirometry has become an indispensable tool in primary care to exclude, diagnose, and monitor chronic respiratory conditions, but the quality of spirometry tests in family practices is a reason for concern. Aim of this study was to investigate whether a combination of e-learning and bimonthly performance feedback would improve spirometry test quality in family practices in the course of 1 year.

METHODS Our study was a cluster trial with 19 family practices allocated to intervention or control conditions through minimization. Intervention consisted of e-learning and bimonthly feedback reports to practice nurses. Control practices received only the joint baseline workshop. Spirometry quality was assessed by independent lung function technicians. Two outcomes were defined, with the difference between rates of tests with 2 acceptable and repeatable blows being the primary outcome and the difference between rates of tests with 2 acceptable blows being the secondary outcome. We used multilevel logistic regression analysis to calculate odds ratios (ORs) for an adequate test in intervention group practices.

RESULTS We analyzed 1,135 tests. Rate of adequate tests was 33% in intervention and 30% in control group practices (OR = 1.3; $P = .605$). Adequacy of tests did not differ between groups but tended to increase with time: OR = 2.2 ($P = .057$) after 3 and OR = 2.0 ($P = .086$) in intervention group practices after 4 feedback reports. When ignoring test repeatability, these differences between the groups were slightly more pronounced: OR = 2.4 ($P = .033$) after 3 and OR = 2.2 ($P = .051$) after 4 feedback reports.

CONCLUSIONS In the course of 1 year, we observed a small and late effect of e-learning and repeated feedback on the quality of spirometry as performed by family practice nurses. This intervention does not seem to compensate the lack of rigorous training and experience in performing spirometry tests in most practices.

Ann Fam Med 2011;9:330-336. doi:10.1370/afm.1258.

INTRODUCTION

Spirometry has become an indispensable tool for primary care professionals to diagnose and monitor chronic respiratory conditions.^{1,2} Recent studies indicate that, when implemented in primary care, spirometry is a valid test³ that leads to increased rates of respiratory diagnoses and may improve disease management in the United States and elsewhere.⁴⁻⁷ Although spirometry can be made available in several ways, from a practical point of view having good-quality tests performed in the practice itself is the preferred mode.⁸

In the Netherlands approximately 62% of family practices own a spirometer (the remaining practices having access to spirometry facilities elsewhere), and the rate of spirometry tests performed in family practices has tripled in the past couple of years.⁹ Recent surveys in the United States indicate that 47% to 75% of family physicians use spirometry.^{10,11}

Lack of spirometry training appears to be an important barrier to the use of spirometry in US family practices.¹⁰⁻¹² Performing good quality spirometry requires proper training and well-standardized procedures,¹³ which may be difficult to achieve in a family practice.¹⁴

Using different sets of criteria, investigators have studied the quality of spirometry tests in family practices and reported rates of adequate tests ranging from only 3% to up to 80%.^{5,14-17} Although there are some indications that training and performance feedback may increase the quality of spirometry in nonlaboratory settings^{18,19}, only 1 controlled intervention trial has been published from a real-life family practice setting.¹⁵ In the study reported in this article, we investigated whether a combined intervention of e-learning and subsequent bimonthly performance feedback after a baseline spirometry workshop would improve test quality in family practices in the course of 1 year.

METHODS

Study Design

The study was a cluster controlled trial, with family practices as clusters for patients in whom 1 or more spirometry tests were performed during a 1-year observation period (Clinicaltrials.gov Protocol Registration System: NCT00962455; <http://www.clinicaltrials.gov>). All practices involved were nonacademic practices that have a working agreement with a local hospital (Elkerliek Hospital, Helmond, The Netherlands) regarding support in spirometry training, test execution, and interpretation. All practices have a PC-based spirometer (SpiroPerfect, Welch Allyn, Delft, The Netherlands). The spirometry tests are electronically submitted by the practices and can be accessed by the hospital's lung function technicians and chest physicians. The Medical Ethics Committee of the Elkerliek Hospital approved the study (file number 07-393).

Recruitment, Sample Size, and Group Allocation

Practices were recruited through a mailing to all family practices that collaborate with the Elkerliek Hospital. Cluster sample size calculation²⁰ showed that 19 practices were needed to demonstrate a 25% difference in the rate of adequate spirometry tests. Assumptions for the power calculation were as follows¹⁴⁻¹⁶: 30% adequate spirometry tests in control practices; 12-month study duration; average of 2 spirometry tests per practice per week; intraclass correlation coefficient (ICC) of 0.15; α level of .05; $1 - \beta$ of 0.80. Although we anticipated a learning curve in intervention effects over time, we did not take this into account in the sample size calculation. After recruitment, practices received a questionnaire to inquire about their spirometry details

and were allocated to either the intervention or usual practice group through computerized minimization, a method of ensuring balance between groups for several prognostic factors, even in small samples.²¹ Allocation was stratified by the average weekly number of spirometry tests in the past year and by practice nurses' years of spirometry experience.

Intervention and Control Conditions

All practice nurses involved in the study attended a 2.5-hour baseline workshop to refresh their spirometry knowledge and skills. Current criteria for adequate spirometry tests¹³ were discussed, and execution of tests was practiced. At the end of the workshop, all nurses were encouraged to continue or increase their usual frequency of spirometry testing. Intervention practice nurses stayed for an additional 30 minutes to (1) receive a copy of the e-learning CD-ROM *Spirometry Fundamentals* (University of Washington, Seattle, Washington), (2) be instructed in how to study the CD-ROM in the next 4 weeks, and (3) discuss details of the upcoming bimonthly feedback reports—including the way these reports should be interpreted and discussed during telephone calls with the lung function technician who had been assigned as their coach. The feedback reports focused on the spirometry tests performed in the past 2- to 2.5-month period and, after the first period, on comparison with previous periods. All practices were instructed that (1) only trained nurses should perform spirometry, (2) all tests performed as a part of routine patient care should be submitted, and (3) each submitted test should consist of the 3 best blows.

The observation period for the trial started immediately after the spirometry workshop (on November 6, 2007) and lasted exactly 12 months. A new feedback period started on the day the feedback reports that were based on the tests from the previous period had been e-mailed to the intervention practices. After electronic submission, the quality of prebronchodilator tests from intervention practices was assessed by one of the Elkerliek Hospital's lung function technicians, using a checklist based on recent recommendations.¹³ Because the lung function technicians could invest a limited number of hours to assess spirometry tests for the study, 25 tests per practice per period was set as the maximum number of tests on which a feedback report could be based.

Spirometry Quality Assessment and Study Outcomes

Three experienced lung function technicians who were not otherwise involved in the study assessed de-identified printouts of all submitted tests. These technicians were instructed on how to score spirometry test quality,¹³ practiced on 10 tests, and discussed their scoring

and experiences. Two technicians volunteered to assess all tests for the trial. Interobserver agreement between these 2 technicians was assessed in a random sample of 80 tests²² and was deemed sufficient (Cohen's $\kappa = 0.79$). In case of disagreement, the third technician assessed the test. Tests were presented to the technicians in random order, and they were blinded to each other's assessments and to information on the practice or study period to which a test belonged.

We only used prebronchodilator tests from patients older than 10 years. The primary outcome for the study was the proportion of tests with 2 or more acceptable blows that were also repeatable for both FEV₁ (forced expiratory volume in 1 second) and FVC (forced vital capacity).^{13,15} An adequate forced blow has a good start, satisfactory duration of exhalation, and is free from artifacts (eg, cough, glottis closure, obstruction of mouth-piece). Because the feedback report focused strongly on acceptability of blows and less so on repeatability between blows, we studied differences in the proportions of tests with 2 or more acceptable blows as a secondary outcome.

Statistical Analysis

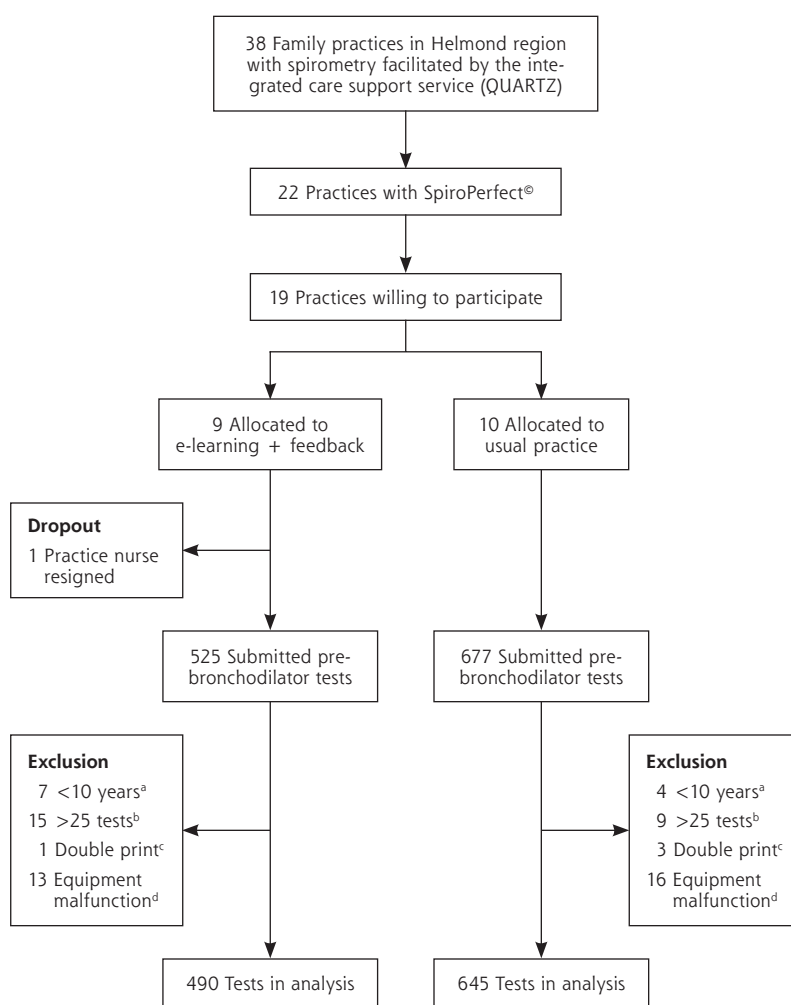
Multilevel logistic regression analysis with family practice as cluster level and a compound symmetry matrix correlation structure was used to test differences in the odds of an adequate spirometry test (GLIMMIX procedure, SAS 9.2, SAS Institute, Cary, North Carolina). Odds ratios (ORs) derived from the logistic regression models express the odds (and 95% confidence interval [CI]) of an adequate test in intervention relative to control practices. The logistic model also included the period in which the test had been performed and covariates related to test quality (ie, age, sex, and severity of airflow obstruction).¹⁶ ICCs were calculated for the primary and secondary outcomes. Statistical tests were 2-sided; $P < .05$ was considered statistically significant.

RESULTS

Practices, Patients, and Spirometry Tests

Figure 1 displays the recruitment and allocation of practices and reasons for excluding spirometry tests from the analysis. One of the 9 intervention practices dropped out before the first feedback report was provided. A total of 490 prebronchodilator spirometry tests from intervention and 645 tests from usual prac-

Figure 1. Flowchart of practice recruitment and selection of spirometry tests for the analysis.



^a In their mutual working agreement, the family practices and the Elkerliek Hospital agreed that patients younger than 16 years should be referred to the hospital when a spirometry test is required. Even so, there were 27 tests from patients in this age-group in our dataset. Because obtaining a good-quality spirometry test in young children requires specific training and skills from nurses, we excluded tests from children younger than 10 years from the analysis.

^b Because the lung function technicians from the Elkerliek Hospital could invest a limited number of hours to assess spirometry tests for the study, 25 tests per practice per period was set as the maximum number on which a feedback report could be based.

^c All tests were color-printed in the pulmonary function laboratory of the Elkerliek Hospital and de-identified by the investigators before the blinded, randomized outcome assessment by the independent lung function technicians. A small number of tests were erroneously printed twice.

^d Tests that had been submitted by family practices in which the independent lung function technicians recognized technical malfunction of the spirometer or spirometry software. These tests were excluded from further analyses because the lung function technicians could not judge the nurse's performance.

Table 1. Characteristics of Family Practices and Patients Tested With Spirometry

Characteristic	Intervention	Control	P Value ^a
Practices			
Number of practices	9	10 ^b	
Practice type, n (%)			
Solo	2 (22)	2 (22)	
Duo	1 (11)	3 (33)	
Group	5 (56)	2 (22)	
Multidisciplinary health care center	1 (11)	2 (22)	
Physicians per practice, mean (SD), n	3.2 (1.6)	2.3 (1.7)	.582
Practice size, patients, mean (SD), n	6,521 (3,428)	4,754 (3,012)	.263
Spirometry test operators per practice, mean (SD), n	2.0 (0.9)	2.3 (0.9)	.461
Spirometry experience, mean (SD), y	5.6 (3.0)	5.0 (3.0)	.673
Spirometry tests per month, mean (SD), n	13.7 (15.0)	16.6 (10.6)	.644
Tests			
Spirometry tests, n	490	645	
Unique patients, n	457	615	
Age, mean (SD), y	53.3 (16.1)	53.9 (17.9)	.550
Male, n (%)	245 (50)	303 (47)	.421 ^c
FEV ₁ , mean (SD), L ^d	2.56 (0.97)	2.55 (0.98)	.829
FEV ₁ , as % of predicted value, mean (SD) ^e	84.5 (23.4)	84.0 (20.8)	.719

FEV₁ = forced expiratory volume in 1 second; FVC = forced vital capacity.

^a From Student t test for independent samples.

^b One baseline questionnaire was not returned in the usual practice group (after 2 reminders).

^c From Pearson χ^2 test.

^d From prebronchodilator spirometry test.

^e Predicted values from the European Community for Steel and Coal (ECSC), 1993.

tics were analyzed. In the intervention group 29 patients contributed 2 tests, and 2 patients contributed 3 tests. The respective numbers of contributed tests for the control group were 26 and 2 patients.

Table 1 displays characteristics of practices and patients. On average, practices employed 2.1 (range = 1-3) spirometry test operators, had 5.3 years (SD = 3.0 years) of experience with spirometry, and performed 15 spirometry tests (SD = 13) per month when beginning the study. Mean age of the patients tested in the practices was 53.6 years (SD = 10.9 years), and 48% were male.

Effects on Primary and Secondary Outcomes

The coaches' records showed that all intervention practice nurses reported to have studied the spirometry CD-ROM shortly after the baseline workshop, and that instructions in the feedback reports to repeat particular CD-ROM modules were followed in 92% of cases. Table 2 shows the rate of test adequacy for the respective

Table 2. Reasons for Lack of Acceptability of Forced Expiratory Blows in the Intervention and Control Groups

Reason	Period 1	Period 2	Period 3	Period 4	Period 5	All Periods
e-Learning + feedback, blow (test), n	306 (102)	330 (113)	261 (87)	273 (91)	291 (97)	1470 (490)
Usual practice, blow (test), n	348 (116)	456 (152)	348 (116)	405 (135)	378 (126)	1935 (645)
Blows with poor start						
e-Learning + feedback, n (%)	21 (20.6)	23 (20.4)	19 (21.8)	16 (17.6)	16 (16.5)	95 (19.4)
Usual practice, n (%)	27 (23.3)	33 (21.7)	28 (24.1)	37 (27.4)	25 (19.8)	150 (23.3)
P value ^a	0.857	0.931	0.933	0.095	0.664	0.474
Blows with artifacts during exhalation						
e-Learning + feedback, n (%)	11 (10.8)	23 (20.4)	18 (20.7)	17 (18.7)	14 (14.4)	83 (16.9)
Usual practice, n (%)	11 (9.5)	17 (11.2)	13 (11.2)	23 (17.0)	14 (11.1)	78 (12.1)
P value ^a	0.644	0.026	0.214	0.987	0.757	0.311
Blows with unsatisfactory exhalation						
Abrupt end						
e-Learning + feedback, n (%)	6 (5.9)	15 (13.3)	10 (11.5)	5 (5.5)	5 (5.2)	41 (8.4)
Usual practice, n (%)	12 (10.3)	9 (5.9)	3 (2.6)	9 (6.7)	5 (4.0)	38 (5.9)
P value ^a	0.190	0.410	0.082	0.710	0.727	0.702
Duration <6 sec or no plateau in volume-time curve						
e-Learning + feedback, n (%)	35 (34.3)	38 (33.6)	30 (34.5)	29 (31.9)	28 (28.9)	160 (32.7)
Usual practice, n (%)	55 (47.5)	63 (41.4)	44 (37.9)	69 (51.1)	49 (38.9)	280 (43.4)
P value ^a	0.180	0.380	0.749	0.171	0.854	0.371

^a From multilevel logistic regression analysis.

quality criteria. Overall, poor start (21.6%) and insufficient duration of blows (38.8%) were the predominant reasons for test inadequacy, but no statistically significant differences between intervention and control group practices were observed. For the primary outcome, the rate of adequate tests in the first 2 months

(ie, after the baseline workshop and initial e-learning in intervention practices, but before the first feedback report) was 30.4% in intervention and 25.0% in control practices. The odds ratio for an adequate test in intervention relative to control practices was 1.3 (95% CI, 0.5-3.2; $P = .605$). For all 5 study periods combined, the rate of adequate tests was 32.9% in the intervention group and 29.8% in the control group (OR=1.2; 95% CI, 0.6-2.5; $P = .663$). Figure 2 (panel A) shows that the adequacy of tests gradually increased over the consecutive study periods in the intervention group relative to the control group. In the fifth and final study period, rates of adequate tests for the primary outcome were 43.3% and 34.1% for intervention and control practices, respectively. The odds ratio of an adequate spirometry test in the intervention group was not statistically significant during the whole test period, but approached the threshold for significance after 3 feedback reports: OR=2.2 (95% CI, 1.0-5.0; $P = .057$) in period 4 and OR=2.0 (95% CI, 0.9-4.5; $P = .086$) in period 5.

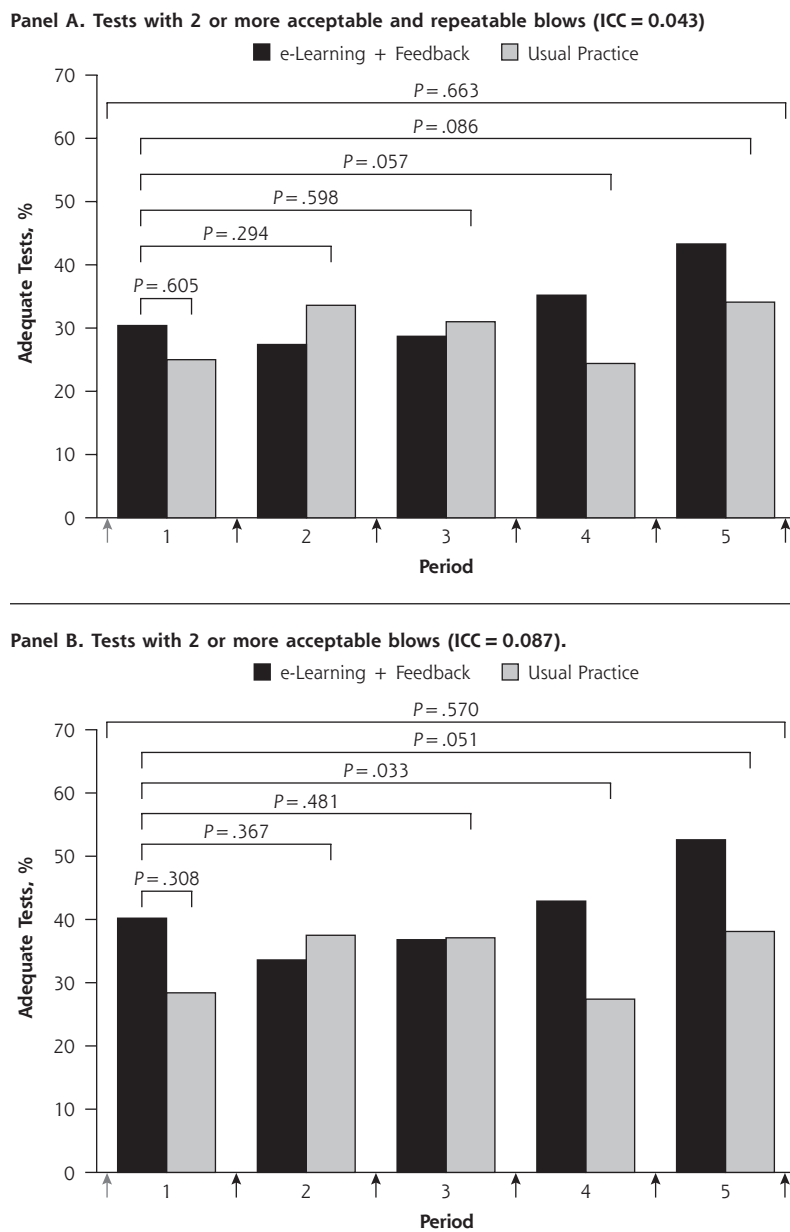
When repeatability between blows was ignored in the definition of test adequacy, the differences between the study groups were more pronounced: OR=2.4 (95% CI, 1.1-5.3; $P = .033$) in period 4 and OR=2.2 (95% CI, 1.0-4.8; $P = .051$) in period 5 (Figure 2, panel B). Supplemental Figure 1 (available at <http://www.annfammed.org/cgi/content/full/9/4/330/DC1>) displays the rate of adequate spirometry tests per practice and per period for the primary outcome.

DISCUSSION

Summary of Main Findings

In a cluster controlled trial we investigated whether a combined intervention of e-learning and bimonthly performance feedback was able to improve the quality of spirometry tests in Dutch family practices that had

Figure 2. Rate of adequate spirometry tests in intervention and usual practice groups during 12 months of follow-up: primary (panel A) and secondary (panel B) outcomes.



ICC = intraclass correlation.

Note: The grey arrow on the far left indicates the baseline spirometry workshop in both groups, and the distribution of *Spirometry Fundamentals* CD-ROMs to the nurses from the intervention practices. The black arrows indicate the timing of the feedback to intervention practices. P values are from multilevel logistic regression analyses with family practices as clusters, controlling for age, sex, and severity of airflow obstruction.

already implemented spirometry within their regular patient care. Although we did not observe an intervention effect for the overall 12-month observation period, a modest increase in the rate of adequate tests emerged after 3 feedback reports. This increase led to a trend toward an approximately 10% higher rate of adequate tests in the final 4 months of the observation period in the intervention group. As in previous studies,^{15,16} not meeting end-of-test criteria was the predominant reason for test inadequacy.

Comparison With Existing Literature

Certified and skilful lung function technicians can be expected to produce adequate spirometry tests in 80% or more of adult patients tested.²³ Although our data show that some family practice nurses may be able to achieve rates of adequate tests in 50% to 75% of tests (Supplemental Figure 1), it is unlikely that most practices will be able to achieve and maintain such a performance level. A recent Australian study showed that delegating spirometry to well-trained and experienced visiting nurses substantially improved spirometry adequacy.¹⁴ Adopting this (or a similar) model may overcome practice nurses' lack of training, experience, and routine—which are essential factors for good-quality spirometry.²⁴ Studies conducted in specialized^{19,25} as well as in primary care^{15,18} settings suggest that performance feedback is able to improve spirometry test quality. Our findings support such an approach, although expectations should be fairly modest, and it may take a rather long breath and more intensive (certified) training programs to achieve satisfactory results.

In our view, obtaining sufficiently reliable and clinically useful spirometry tests—not necessarily perfect tests—is what family practices should be striving for. The impact is of inadequate spirometry tests on patients' diagnoses and management in family practice is currently unclear.

Strengths and Limitations

A particular strength of our study was that it was undertaken in a real-life setting in which a group of family practices had implemented spirometry for several years. Compared with national figures, our group of practices was representative for the Netherlands in terms of spirometry experience,⁹ but smaller (ie, single-handed and duo) practices were underrepresented.²⁶ All practices involved in the evaluation participated in the working agreement with the local hospital, which may limit the generalizability of our findings. Other strengths are the high participation rate, the individualized feedback, the rigorous method of outcome assessment using multiple blinded experts, the clustered trial design and multilevel analyses, and the stratification by

spirometry experience of practice nurses. A limitation was the 12-month duration of the study. Although 12 months would seem like a sufficient observation period, the intervention appeared to start having effect after 3 feedback reports. A few additional months of follow-up might have shown either a continuing upward trend, a plateau, or a regression of intervention effects.

During the course of 1 year we found a rather small and late effect of baseline e-learning and repeated performance feedback on the quality of spirometry tests by practice nurses who perform spirometry as a part of regular patient care in family practices. This intervention does not seem to be able to fully compensate for their lack of rigorous training and experience in performing spirometry tests. Other models to provide family practices with good-quality spirometry should be explored.

To read or post commentaries in response to this article, see it online at <http://www.annfammed.org/cgi/content/full/9/4/330>.

Key words: Spirometry; respiratory function tests; family practice; primary health care; quality assurance, health care; randomized controlled trial; multicenter study; education; pulmonary disease, chronic obstructive; asthma

Submitted September 3, 2010; submitted, revised, January 5, 2011; accepted January 18, 2011.

Funding support: This study was funded by QUARTZ, a regional, integrated care support service; the Elkerliek Hospital; and the Department of Primary and Community Care of the Radboud University Nijmegen Medical Centre, the Netherlands.

An oral presentation at the International Primary Care Respiratory Group (IPCRG) Conference in Toronto, Canada, 2-5 June 2010.

Acknowledgements: The authors wish to express their gratitude to the family practice nurses, family practitioners, chest physicians, and patients who contributed to our study. The authors owe much gratitude to Cyril Welten and Marloes Beekwilder, the lung function technicians from the Elkerliek Hospital in Helmond, who assessed the spirometry tests from the intervention practices and coached the practice nurses from these practices. Without their enthusiasm and commitment we would not have been able to conduct this study. The authors very much appreciate the time and effort put in the study by Oda van Heeswijk, André Hijmans and Kees Lamers, the lung function technicians from the Radboud University Nijmegen Medical Centre, who judged the spirometry tests for the outcome assessment. The authors also appreciate the contributions to the preparatory and supportive work on the study by Janny Wildschut, Wilfried van Huisseling, Rasha Putros, Nicol Orbon, Daan Schellekens, and Hans Vlek.

References

1. Levy ML, Fletcher M, Price DB, Hausen T, Halbert RJ, Yawn BP. International Primary Care Respiratory Group (IPCRG) Guidelines: diagnosis of respiratory diseases in primary care. *Prim Care Respir J*. 2006;15(1):20-34.
2. Bellamy D, Bouchard J, Henrichsen S, et al. International Primary Care Respiratory Group (IPCRG) Guidelines: management of chronic obstructive pulmonary disease (COPD). *Prim Care Respir J*. 2006; 15(1):48-57.

3. Schneider A, Gindner L, Tilemann L, et al. Diagnostic accuracy of spirometry in primary care. *BMC Pulm Med*. 2009;9:31.
4. Walker PP, Mitchell P, Diamantea F, Warburton CJ, Davies L. Effect of primary-care spirometry on the diagnosis and management of COPD. *Eur Respir J*. 2006;28(5):945-952.
5. Yawn BP, Enright PL, Lemanske RF Jr, et al. Spirometry can be done in family physicians' offices and alters clinical decisions in management of asthma and COPD. *Chest*. 2007;132(4):1162-1168.
6. Dales RE, Vandemheen KL, Clinch J, Aaron SD. Spirometry in the primary care setting: influence on clinical diagnosis and management of airflow obstruction. *Chest*. 2005;128(4):2443-2447.
7. Lusuardi M, De Benedetto F, Paggiaro P, et al. A randomized controlled trial on office spirometry in asthma and COPD in standard general practice: data from spirometry in asthma and COPD: a comparative evaluation Italian study. *Chest*. 2006;129(4):844-852.
8. Schermer T, Eaton T, Pauwels R, van Weel C. Spirometry in primary care: is it good enough to face demands like World COPD Day? *Eur Respir J*. 2003;22(5):725-727.
9. Schellekens D, Poels P, Pellegrino A, Cretier R, Smeele I, Schermer T. [Spirometry in Dutch family practice. Results from a national survey.] *Huisarts Wet*. 2008;51(9):434-439.
10. Blain EA, Craig TJ. The use of spirometry in a primary care setting. *Int J Gen Med*. 2009;2:183-186.
11. Dombkowski KJ, Hassan F, Wasilevich EA, Clark SJ. Spirometry use among pediatric primary care physicians. *Pediatrics*. 2010;126(4):682-687.
12. O'Dowd LC, Fife D, Tenhave T, Panettieri RA Jr. Attitudes of physicians toward objective measures of airway function in asthma. *Am J Med*. 2003;114(5):391-396.
13. Levy ML, Quanjer PH, Booker R, Cooper BG, Holmes S, Small I. Diagnostic spirometry in primary care: Proposed standards for general practice compliant with American Thoracic Society and European Respiratory Society recommendations: a General Practice Airways Group (GPIAG) document, in association with the Association for Respiratory Technology & Physiology (ARTP) and Education for Health. *Prim Care Respir J*. 2009;18(3):130-147.
14. Walters JA, Hansen EC, Johns DP, Walters EH, Blizzard L, Wood-Baker R. A mixed methods study to compare models of spirometry delivery in primary care for patients at risk of Chronic Obstructive Pulmonary Disease. *Thorax*. 2008;63(5):408-414.
15. Eaton T, Withy S, Garrett JE, Mercer J, Whitlock RM, Rea HH. Spirometry in primary care practice: the importance of quality assurance and the impact of spirometry workshops. *Chest*. 1999;116(2):416-423.
16. Schermer TR, Crockett AJ, Poels PJ, et al. Quality of routine spirometry tests in Dutch general practices. *Br J Gen Pract*. 2009;59(569):e376-e382.
17. Tuomisto L, Jarvinen V, Laitinen J, Erhola M, Kaila M, Brander P. Asthma Programme in Finland: the quality of primary care spirometry is good. *Prim Care Respir J*. 2008;17(4):226-231.
18. Burton MA, Burton DL, Simpson MD, Gissing PM, Bowman SL. Respiratory function testing: the impact of respiratory scientists on the training and support of primary health care providers. *Respirology*. 2004;9(2):260-264.
19. Upton MN, Ferrell C, Bidwell C, et al. Improving the quality of spirometry in an epidemiological study: The Renfrew-Paisley (Mid-span) family study. *Public Health*. 2000;114(5):353-360.
20. Campbell MK, Thomson S, Ramsay CR, MacLennan GS, Grimshaw JM. Sample size calculator for cluster randomized trials. *Comput Biol Med*. 2004;34(2):113-125.
21. Altman DG, Bland JM. Treatment allocation by minimisation. *BMJ*. 2005;330(7495):843.
22. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17(1):101-110.
23. Enright PL, Beck KC, Sherrill DL. Repeatability of spirometry in 18,000 adult patients. *Am J Respir Crit Care Med*. 2004;169(2):235-238.
24. Malmstrom K, Peszek I, Al Botto, Lu S, Enright PL, Reiss TF. Quality assurance of asthma clinical trials. *Control Clin Trials*. 2002;23(2):143-156.
25. Enright PL, Johnson LR, Connett JE, Voelker H, Buist AS. Spirometry in the Lung Health Study. 1. Methods and quality control. *Am Rev Respir Dis*. 1991;143(6):1215-1223.
26. Poels PJ, Schermer TR, Jacobs A, et al. Variation in spirometry utilization between trained general practitioners in practices equipped with a spirometer. *Scand J Prim Health Care*. 2006;24(2):81-87.