Impact of a spirometry expert system on general practitioners' decision making

P.J.P. Poels*, T.R.J. Schermer*, D.P.A. Schellekens*, R.P. Akkermans*,
P.F. de Vries Robbé#, A. Kaplan¶, B.J.A.M. Bottema+ and C. van Weel*

ABSTRACT: The present study assessed the impact of computerised spirometry interpretation expert support on the diagnostic achievements of general practitioners (GPs), and on GPs' decision making in diagnosing chronic respiratory disease.

A cluster-randomised controlled trial was performed in 78 GPs who each completed 10 standardised paper case descriptions. Intervention consisted of support for GPs' spirometry interpretation either by an expert system (expert support group) or by sham information (control group). Agreement of GPs' diagnoses was compared with an expert panel judgement, which served as the primary outcome. Secondary outcomes were: additional diagnostic test rates; width of differential diagnosis; certainty of diagnosis; estimated severity of disease; referral rate; and medication or nonmedication changes. Effects were expressed as odds ratios (ORs) with 95% confidence intervals (CIs).

There were no differences between the expert support and control groups in the agreement between GPs and expert panel diagnosis of chronic obstructive pulmonary disease (OR (95% CI) 1.08 (0.70–1.66)), asthma (1.13 (0.70–1.80)), and absence of respiratory disease (1.32 (0.61–2.86)). A higher rate of additional diagnostic tests was observed in the expert support group (2.5 (1.17–5.35)).

Computerised spirometry expert support had no detectable benefit on general practitioners' diagnostic achievements and the decision-making process when diagnosing chronic respiratory disease.

KEYWORDS: Computer-assisted diagnosis, expert systems, family practice, spirometry

Although all major chronic obstructive pulmonary disease (COPD) guidelines stress the central role of spirometry in diagnosing and managing chronic respiratory disease [1, 2], this does not guarantee that general practitioners (GPs) will consequently use spirometry in the care of their patients with respiratory symptoms [3, 4].

Most common barriers that impede utilisation of spirometry in general practice are: the absence of properly trained staff [5]; the lack of time and practice support to fit spirometry into the daily practice routine [6]; the absence of a spirometer in the practice [7]; and GPs' lack of confidence in the ability to interpret the test results [8, 9]. A recent survey [4] showed that one third of Australian GPs interpreted less than one spirometry test per week. Due to this low prevalence of test interpretations, it seems difficult for GPs to become experts in this area.

The present authors have previously demonstrated the influence of spirometry on GPs' diagnostic achievements and management decisions in a nonrandomised simulation study [10]. Other recent nonrandomised studies [11, 12] confirm that spirometry increases diagnostic rates of chronic respiratory disease and may lead to management changes in a general practice population. However, an absolute prerequisite for the use of spirometry is the validity (or reliability) of spirometric tests. In a previous study with patients with COPD, SCHERMER et al. [13] observed that the most relevant indices, as measured by trained general practice staff, were comparable with those measured in pulmonary function laboratories.

Therefore, once GPs have had initial spirometry training and spirometry equipment and test validity are adequate, the next step to improve implementation of spirometry in general practice

This article has supplementary material accessible from www.erj.ersjournals.com
For editorial comments see page 8.

is to arrange for the possibility to receive continuous advice and support for test interpretation [14]. This could be carried out by means of a diagnostic computerised clinical decision support system [15, 16]. While there is already such an expert support system available on the market [17] and GPs welcome such type of support [18], empirical studies on the effects of ongoing expert support on the interpretative capacity and self-confidence of GPs are warranted.

The objective of the present study was to assess the impact of expert support for the interpretation of spirometry tests on GPs' diagnostic achievements and decision-making processes when diagnosing chronic respiratory disease.

METHODS
Study design
The study was a simulated cluster-randomised trial of GPs' diagnostic acuity of chronic respiratory disease in a process of diagnostic assessment of 10 standardised cases, with an expert system support. A diagnosis of the cases by the expert panel served as the gold standard. Differences in GPs' diagnostic achievements and decision-making processes were compared both between the study groups and within groups.

Ethical approval
The present study was approved by the Medical Ethics Review Board of the academic hospital Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.

Participants
GPs from the catchment area of the Radboud University Nijmegen Medical Centre and from a specific general practice network of the present authors' department at this hospital [19] were invited to participate by postal mailing.

Intervention
GPs were randomly allocated to one of the following two groups: 1) the computerised spirometry expert interpretation support group; and 2) the control group. GPs in the expert support group received the spirometry test results, the flow–volume curve, and the graphical interpretation and textual interpretative notes. GPs in the control group received the spirometry test results, and the flow–volume and volume–time curves (fig. 1).

The spirometry expert system (SpidaXpert®; Micro Medical Ltd, Rochester, UK) [17] contains a diagnostic algorithm based on pre- and post-bronchodilator forced expiratory volume in one second (FEV_1)/forced vital capacity (FVC) and FEV_1 values and the accompanying age, sex and ethnicity-specific predicted values. The expert interpretation module in SpidaXpert® had been developed with funding of the Netherlands Asthma Foundation by a group of independent experts [17]. The spirometry interpretation is presented as coloured bars that indicate levels of FEV_1/FVC and FEV_1, and compares the values before and after bronchodilatation. The graphical representation is further elucidated by a textual interpretation, which provides information on and suggestions for additional diagnostic testing and treatment options.

GPs in the control group received the volume–time curve as sham information. Sham information was introduced in the control group to be able to compare GPs reassessment of a

diagnosis in the control group in the same way as in the expert support group. Sham information has, in fact, a placebo effect, as no new data was being presented to these GPs; earlier data (*i.e.* the flow–volume curve) was presented in each case again but in another way, *i.e.* the volume–time curve. Although it is clearly important to evaluate the quality of forced expiratory manoeuvres, *i.e.* end-of-test criteria [20], the volume–time curve does not add relevant new information from a diagnostic point of view to the information provided by the flow–volume curve and the numerical test results. Prior to the study, participants were informed that they would receive additional information on spirometry and were asked to reconsider their diagnosis. No further specification was given of the nature or the background of that information.

Standardised case descriptions and gold standard
Based on the present authors' experiences in a previous study [10], it was known beforehand that GPs are quite able to diagnose common respiratory disease patterns, whereas rare pathologies and inadequate test results are more difficult for them to recognise. Furthermore, the challenge to differentiate COPD from other conditions that result in respiratory symptoms (*e.g.* heart failure, asthma) grows with the age of the patient. This was the reason for including case descriptions of adult patients only, with a special focus on the 50–60-yr-old age group. This category reflects daily practice patterns in primary care. The case descriptions, in which a GP would use spirometry as a diagnostic test, were as follows: COPD (classified as Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage I (n=1), stage II (n=1) and stage III (n=2)) [2]; asthma (n=2); allergic asthma (n=1); lung fibrosis (n=1); no respiratory disease (n=1); incorrect test manoeuvre (n=1); and exercise-induced asthma (n=1; see supplementary material for example case).

At inclusion, a research assistant visited the participating GPs in their practice. During a 90-min audiotaped session, an example case and 10 standardised cases were presented on a laptop computer using PowerPoint slides. GPs worked through the cases in a random order. GPs first practised on one separate example case to become familiar with case structure. For each case, a concise medical history, the results of physical examination and the medication were presented to the GP first. Subsequently, absolute predicted pre- and post-bronchodilator spirometry test results (including FEV_1, FVC, FEV_1/FVC and flow–volume curves) were provided. GPs were asked to consider their diagnosis and management before the upcoming intervention. Next, GPs received additional information next to the spirometry test results: either the graphical representation of FEV_1, FEV_1/FVC together with interpretative notes (expert support group) or the volume–time curve (control group). Again GPs were asked to reconsider their diagnosis and management after the intervention. An example of the case structure is depicted in figure 2. Due to time limitations, the present authors requested only for specific medication and nonmedication changes after the intervention in cases with already diagnosed respiratory disease (six out of 10 cases).

Before their use in the study, the cases were judged by an expert panel consisting of two chest physicians, a GP (P.J.P. Poels) with specific expertise in spirometry and a health
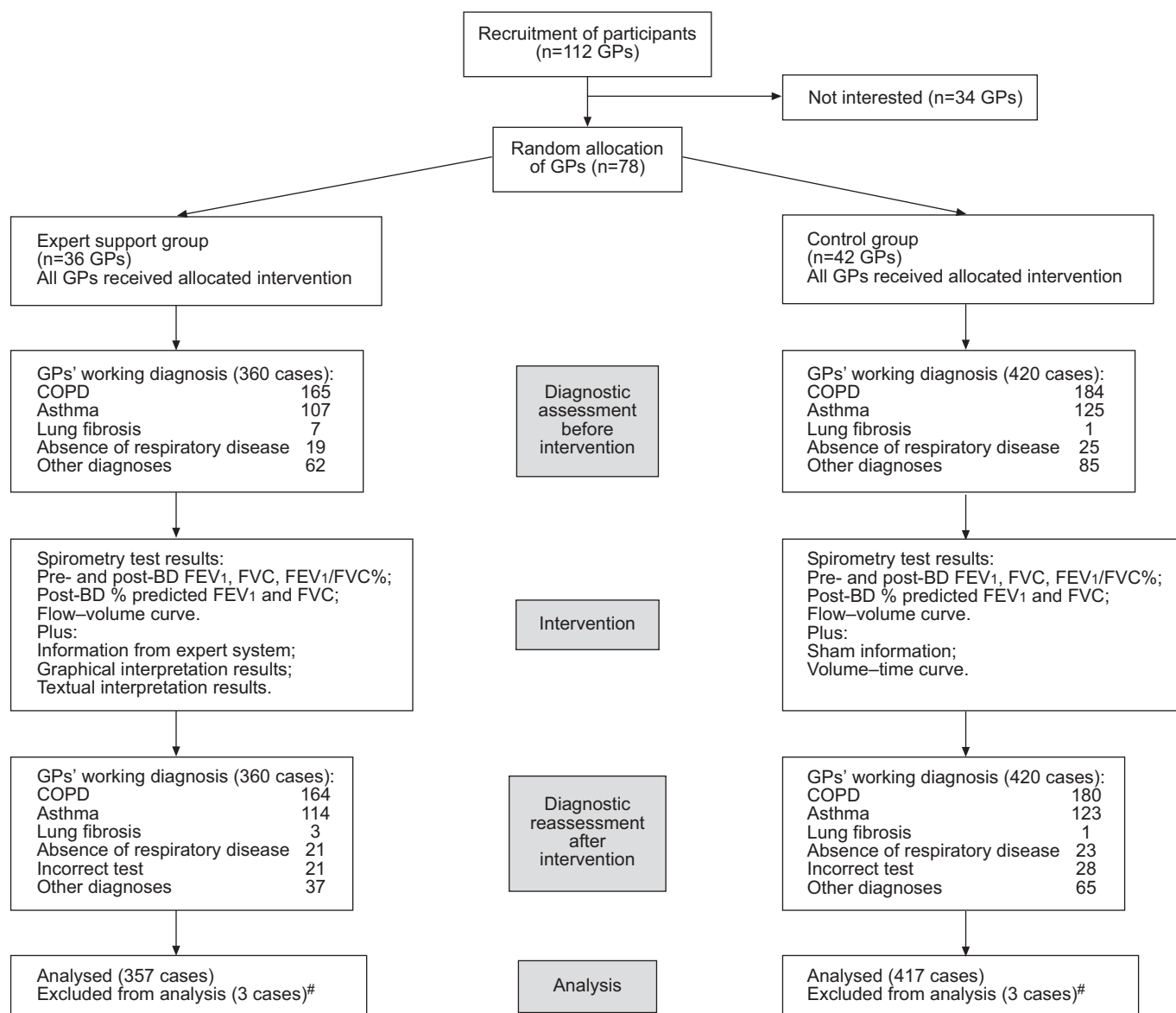
Recruitment of participants
(n=112 GPs)

Not interested (n=34 GPs)

Random allocation
of GPs (n=78)

Expert support group
(n=36 GPs)
All GPs received allocated intervention

Control group
(n=42 GPs)
All GPs received allocated intervention

GPs' working diagnosis (360 cases):
COPD   165
Asthma   107
Lung fibrosis   7
Absence of respiratory disease   19
Other diagnoses   62

Diagnostic
assessment
before
intervention

GPs' working diagnosis (420 cases):
COPD   184
Asthma   125
Lung fibrosis   1
Absence of respiratory disease   25
Other diagnoses   85

Spirometry test results:
Pre- and post-BD FEV_1, FVC, FEV_1/FVC%;
Post-BD % predicted FEV_1 and FVC;
Flow–volume curve.
Plus:
Information from expert system;
Graphical interpretation results;
Textual interpretation results.

Intervention

Spirometry test results:
Pre- and post-BD FEV_1, FVC, FEV_1/FVC%;
Post-BD % predicted FEV_1 and FVC;
Flow–volume curve.
Plus:
Sham information;
Volume–time curve.

GPs' working diagnosis (360 cases):
COPD   164
Asthma   114
Lung fibrosis   3
Absence of respiratory disease   21
Incorrect test   21
Other diagnoses   37

Diagnostic
reassessment
after
intervention

GPs' working diagnosis (420 cases):
COPD   180
Asthma   123
Lung fibrosis   1
Absence of respiratory disease   23
Incorrect test   28
Other diagnoses   65

Analysed (357 cases)
Excluded from analysis (3 cases)[#]

Analysis

Analysed (417 cases)
Excluded from analysis (3 cases)[#]

FIGURE 1. Participants to the present study and intervention. GP: general practitioner; COPD: chronic obstructive pulmonary disease; BD: bronchodilatation; FEV_1: forced expiratory volume in one second; FVC: forced vital capacity. [#]: the first six GPs used an example case with an expert panel's diagnosis of ''absence of respiratory disease'' and a test case of exercise asthma. For the other 72 consecutive GPs, the case set was switched between these two cases: the case of absence of respiratory disease was included for them in the final case set. Therefore, information was not available from the first six GPs about the case of absence of respiratory disease, equally divided among expert support (n=3 GPs) and control group (n=3GPs).

scientist (T.R.J. Schermer). The panel consensus diagnoses served as the gold standard in the subsequent evaluation of GPs' diagnostic achievements.

The whole approach was piloted in four GPs before the start of the study. Shortly after the first six study visits, the case set was adjusted by switching the example case with a case out of the actual set. As a result, no data of the new introduced case were available for those first six GPs (equally divided over the two groups).

Primary and secondary outcome measures

The difference between the percentage agreement of the cases' diagnoses between GPs and expert panel judgement before

and after interpretation of spirometry served as the primary outcome. Diagnoses were directed to the following five outcome categories: 1) COPD; 2) asthma; 3) rare respiratory pathology (lung fibrosis); 4) absence of respiratory disease; and 5) incorrect test manoeuvre.

Six predefined secondary outcome measures were assessed using indicators that show the impact of the expert system intervention on the GPs decision-making processes, as follows: 1) probability of ordering additional diagnostic tests (yes/no); 2) width of the differential diagnoses (i.e. the working diagnosis plus the number of alternative diagnoses considered by the GP); 3) a GP's certainty of the working diagnosis (self-scored 0–10, with 0=uncertain and 10=certain); 4) a GP's
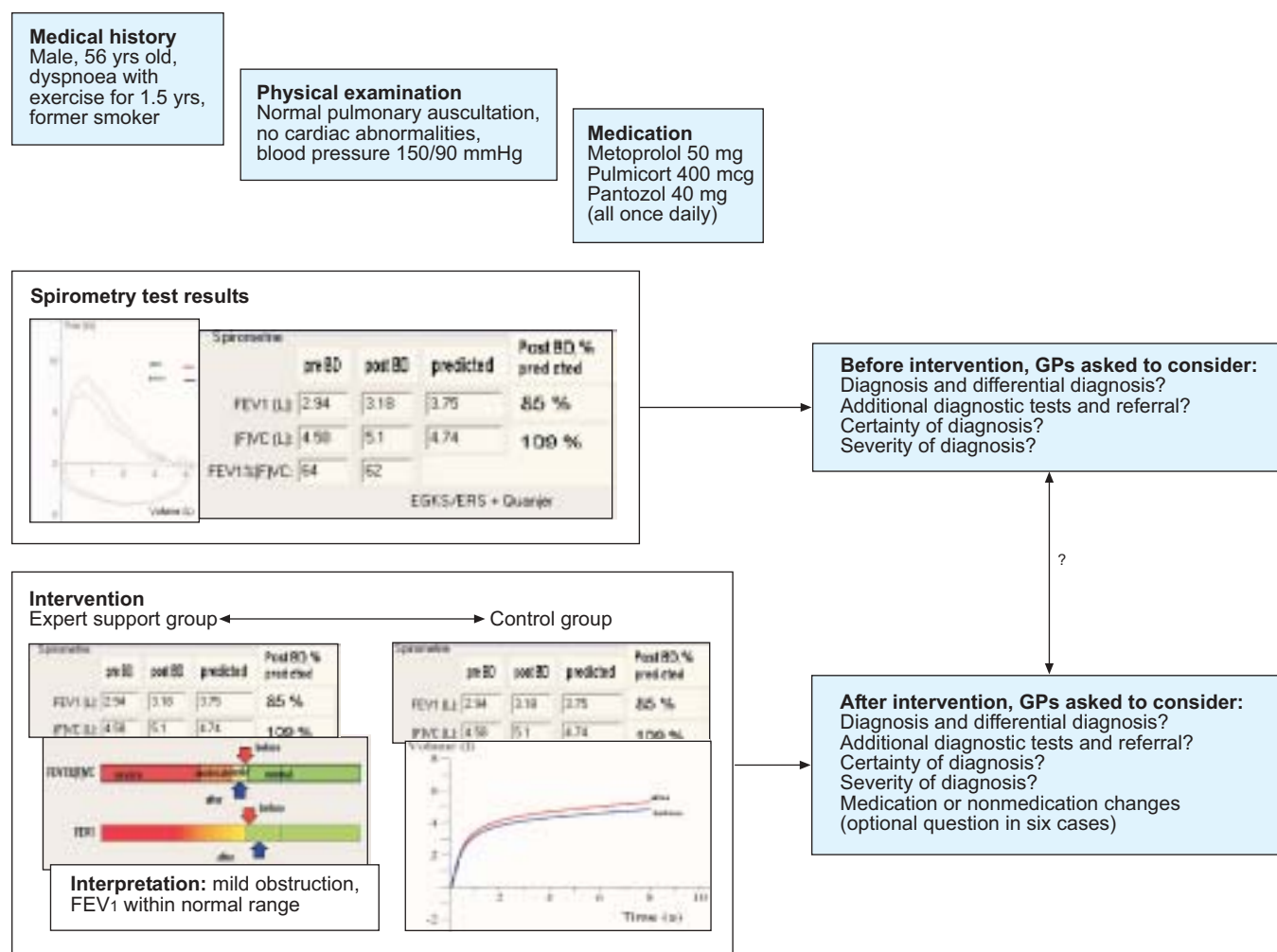
FIGURE 2. Schematic representation of a case structure, detailing information presented to the general practitioner (GP) and points of consideration. BD: bronchodilatation; FEV1: forced expiratory volume in one second.

perception of severity of the working diagnosis (self-scored 0–10, with 0=no severe disease and 10=severe disease); 5) probability of referral to secondary care (yes/no); and 6) probability of medication and nonmedication changes. Medication change included: stopping or lowering treatment with inhaled corticosteroids or bronchodilators; the commencement of bronchodilator, inhaled or oral corticosteroid treatment; or combination drug treatment. Nonmedication included giving smoking cessation advice.

Sample size
Calculation of the sample size was based on an estimated relevant proportion of correctly interpreted cases after spirometry expert support of 25% compared with no expert support. Assuming a correctly interpreted proportion of cases without support of 50% [5], $\alpha=0.05$, a power of 80% and an intra-cluster correlation $r=0.18$, 31 GPs were required in each randomisation group. To allow for dropouts and subgroup analyses, the aim was to include $\geqslant70$ GPs.

Randomisation
The research assistant used restricted randomisation (minimisation) with a computer program on a laptop computer

using the following three stratification factors: 1) a GP's prior experience with the specific computerised spirometry interpretation support package (yes/no); 2) the average number of spirometry tests a GP reported to interpret per week; and 3) a GP's experience (in years) with spirometry. The researchers and the statistician (R.P. Akkermans) were blinded while assessing and reporting all outcomes.

Statistical analysis
Agreement between GPs' and expert panel judgement was expressed as percentages with 95% confidence intervals (CIs). Multilevel regression logistic modelling was used to account for the intracluster correlation induced by the fact that each GP assessed more than one case, and the fact that the same cases were applied repeatedly in different GPs. Multilevel logic analyses were performed for dichotomous variables and multilevel regression analyses for continuous variables. Odds ratios (ORs) with 95% CIs were calculated to evaluate differences in percentages of agreement before and after the intervention with the expert judgement between the study groups. Sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively), and the diagnostic OR (DOR) [21] with 95%

CIs were calculated for GP judgements of COPD, asthma, rare respiratory pathology and no respiratory disease after the intervention. ORs with 95% CIs were also used to evaluate differences in indicators GPs' decision-making process.

To detect possible effect modifications before intervention, subgroup analyses were performed for a GP's prior experience with spirometry, a GP's prior experience with expert support and a GP's number of interpreted spirometry tests per week.

RESULTS

Baseline characteristics of GPs

Between January and October 2006, 78 GPs were enrolled in the present study; 36 were allocated to the expert support group and 42 to the control group (fig. 1). All GPs completed the study. Relevant characteristics at baseline were similar between the two groups (table 1).

Primary outcome: diagnostic achievements by GPs

GPs assessed a total of 774 cases, 357 cases from the expert support group and 417 cases from the control group. There was no difference between the expert support and control group in agreement on judgement between GPs and the expert panel for presence of COPD, asthma, absence of respiratory disease and incorrect test manoeuvre after intervention (table 2). GPs' agreement with the expert panel for all cases, except the incorrect test manoeuvre case, was 66.0 (expert support) *versus* 65.9% (control) before intervention and 68.5 (expert support) *versus* 63.5% (control) after intervention.

Although the DORs in the expert support group were consistently higher than in the control group, no significant differences were found between the groups (table 3). GPs did not recognise an incorrect test manoeuvre in 28.6% (in both expert support and control groups) of cases. The highest NPVs were found for cases with the conditions of asthma and absence of respiratory disease.

Secondary outcomes: indicators of GPs' decision-making process

GPs in the expert support group ordered slightly more additional diagnostic tests compared with the control group (OR (95% CI) 2.5 (1.2–5.4); table 4). There were no significant differences between the two groups for other secondary outcome measures. There were also no specific changes (start, stop or lower) in medication (bronchodilators, inhaled steroids or nonpulmonary drugs) between the study groups.

Subgroup analyses

Neither a GP's experience with spirometry (OR (95% CI) 1.02 (0.97–1.06)), nor a GP's prior experience with expert support (0.97 (0.72–1.31)) or a GP's number of interpreted spirometry tests per week (1.02 (0.84–1.23)) was associated with the effectiveness of expert support, as their agreement with the expert panel was not different before intervention. If GPs interpreted more spirometry tests per week and had prior experience of expert support, the probability of agreement with the expert panel before intervention increased; however, this probability decreased if GPs had no prior experience with expert support (interaction effect p=0.02).

DISCUSSION

Statement of principal findings

Computerised spirometry expert support for the interpretation of spirometry tests by GPs had no detectable benefit over sham information on GPs' diagnostic achievements of chronic respiratory disease. Overall, expert support did not influence GPs' decision-making processes.

Strengths of the study

The present study is the first diagnostic study to assess the impact of a commercially available computerised expert support system for spirometry in a randomised simulation study in primary care. The study used standardised patients, which meant that all participants were faced with the same diagnostic challenges. This could only be achieved in an *in vitro* design, as the mix of practice patients in real life would make it difficult to capture the necessary variation in diagnostic challenges.

The standardised complex and original method that was used to assess the impact of expert support in the present study has been used before in a nonrandomised design [10]. Based on previous information [10], the present authors were able to create a balanced mixture of cases relevant for GPs. The confirmative role of spirometry was more strongly

TABLE 1	Baseline characteristics of all randomised general practitioners (GPs)		
		Expert support group	Control group
GPs n		36	42
Type of practice			
Single handed		2 (5)	4 (10)
Duo		9 (25)	9 (21)
Group (≥3 GPs)		15 (42)	19 (45)
Multidisciplinary healthcare centre		10 (28)	10 (24)
Male %		64	57
GP's experience with spirometry in yrs		5.5±4.3	5.3±3.3
Spirometry results interpreted per week		1.4±0.8	1.4±0.7
Prior experience with expert support % yes		47	36

Data are presented as n (%) or mean±SD, unless otherwise indicated.

TABLE 2	Agreement on case diagnoses between general practitioners (GPs) and expert panel judgement before and after intervention					
GP diagnosis	**Expert support group[#]**		**Control group[¶]**		**Expert panel judgement**	**OR[+] (95% CI)**
	Before	**After**	**Before**	**After**		
Presence of						
COPD	32.5	32.5	32.4	30.7	40	1.08 (0.70–1.66)
Asthma	23.5	25.2	23.5	23.0	30	1.13 (0.70–1.80)
Rare respiratory pathology	0.6	0.3	0.0	0.0	10	NA
Absence of respiratory disease	2.8	3.6	3.4	4.4	10	1.32 (0.61–2.86)
Incorrect test manoeuvre	NA	5.9	NA	6.71	10	0.87 (0.48–1.56)

Data are presented as %, unless otherwise indicated. OR: odds ratio; CI: confidence interval; COPD: chronic obstructive pulmonary disease; NA: not available. [#]: n=357; [¶]: n=417; [+]: ORs express the difference in GPs' judgement before and after intervention, *i.e.* expert support compared with control.

focussed on than the exclusive role of spirometry in primary care. To avoid bias, cases were presented in a (computer-generated) random order and analyses were performed blinded for both the investigators and the statistician.

Subgroup analyses showed no difference in baseline diagnostic achievements of GPs with prior experience of spirometry or of the expert support system used, and on the number of spirometry tests a GP interpreted per week. Therefore, the external validity seems quite good, given the fact that the participants were not specifically interested in spirometry.

Possible limitations
The present trial has some limitations. In a diagnostic assessment of chronic respiratory disease, a GP's consideration to perform spirometry in case of an intermediate prior probability of disease is a great diagnostic step [21]. This step was already foreseen in the present study design. The next step of diagnostic refinement does not seem to influence extensively the posterior probability. In the present study, the diagnostic achievements of GPs in both groups were high (prior probability of a correct diagnosis was ~66%). Overall, only 4.3% of initial diagnoses changed after intervention. As the posterior probability in both groups was nearly the same as the prior probability, the role for expert support to change diagnosis and management was very small. Furthermore, the diagnostic achievements of the GPs exceeded the present authors' assumptions in the power calculation (50% correct diagnoses without expert support). It is probable that instruction and support for these GPs had not been effective, as these

TABLE 3	Sensitivity, specificity, predictive values and diagnostic odds ratios (ORs) for general practitioners' judgement after intervention			
	COPD	**Asthma**	**Rare respiratory pathology**	**Absence of respiratory disease**
Expert support group				
Sensitivity	80.6	83.3	2.8	39.4
Specificity	77.5	90.4	99.4	97.5
PPV	70.7	78.9	33.3	61.9
NPV	85.5	92.6	90.1	94.0
Diagnostic OR	14.2	46.9	4.6	25.7
95% CI	8.46–23.98	24.35–90.23	0.59–35.90	9.74–67.71
Control group				
Sensitivity	76.2	76.2	0.0	35.9
Specificity	79.1	90.7	99.7	97.6
PPV	71.1	78.0	0.0	60.9
NPV	83.1	89.8	89.9	93.7
Diagnostic OR	12.1	31.3	NA	23.0
95% CI	7.60–19.34	17.73–55.22	0.0–34.79	9.22–57.17
p-value	0.65	0.36	NA	0.87

Data are presented as %, unless otherwise indicated. COPD: chronic obstructive pulmonary disease; PPV: positive predictive value; NPV: negative predictive value; CI: confidence interval; NA: not available.

TABLE 4	Impact of the intervention on six indicators of general practitioners' (GPs) decision-making process				
Indicators	**Expert support group[#]**		**Control group[¶]**		**OR[+] (95% CI)**
	Before	**After**	**Before**	**After**	
Additional diagnostic tests	70.2	5.2	75.3	3.0	2.5 (1.17–5.35)*
Radiographic imaging	48.7	0.9	51.3	0.8	1.27 (0.25–6.41)
Blood tests	29.2	0.9	39.1	1.3	1.0 (0.65–1.55)
Lung function	13.2	2.3	15.2	0.3	1.0 (0.59–1.70)
Prednisone course	11.2	2.6	8.9	1.0	1.0 (0.61–1.63)
Electrocardiography	3.6	0.3	6.0	0.2	1.0 (0.55–1.80)
Other[§]	1.1	0.0	0.7	0.0	NA
Width of differential diagnoses	2.2 ± 1.0	1.7 ± 0.9	2.3 ± 1.0	1.7 ± 1.0	1.0 (0.81–1.19)
Certainty of diagnosis	6.8 ± 2.0	7.1 ± 1.9	7.3 ± 1.9	7.3 ± 1.8	1.0 (0.57–1.43)
Perception of severity of diagnosis	6.0 ± 2.2	5.9 ± 2.3	6.3 ± 2.1	6.3 ± 2.1	1.0 (0.57–1.42)
Referral rate	18.6	1.7	17.8	2.5	1.0 (0.60–1.66)
Medication and nonmedication changes[ƒ]		76.2		69.1	1.44 (0.80–2.59)
% yes					
Stop or lower medication					
Inhaled corticosteroids	NA	10.2	NA	6.3	1.67 (0.85–3.28)
Bronchodilators	NA	0.9	NA	0.8	1.17 (0.16–8.40)
Start medication					
Short-acting bronchodilators	NA	26.9	NA	22.2	1.27 (0.73–2.20)
Long-acting bronchodilators	NA	15.3	NA	15.5	1.01 (0.46–2.18)
Inhaled corticosteroids	NA	31.0	NA	31.0	1.00 (0.68–1.49)
Oral corticosteroids	NA	6.5	NA	5.6	1.34 (0.44–4.10)
Combinational drug	NA	3.7	NA	2.8	1.27 (0.26–6.19)
Nonmedication changes					
Smoking cessation advice	NA	30.1	NA	33.7	0.85 (0.57–1.26)

Data are presented as % or mean±SD, unless otherwise indicated. OR: odds ratio; CI: confidence interval; NA: not available. [#]: n=357; [¶]: n=417; [+]: ORs express the difference in an indicator of the GPs' decision-making process before and after the intervention, *i.e.* expert support compared with control; [§]: includes urine test, gastroscopy, ergometry, blood pressure, temperature and oxygen saturation; [ƒ]: this information was available for six out of 10 cases (expert support n=216, control group n=252 cases). *: $p < 0.05$.

GPs could already be considered experts due to prior participation in other studies or postgraduate spirometry training programmes from the present authors' department. Therefore, the expert system had hardly additional value and could be considered a ''sort of luxury appendix'' for these GPs.

A large within-group difference was found for ordering additional diagnostic tests, which may be an effect of the study design: GPs barely reassessed their diagnostics after intervention, because they expected the results of their diagnostics to have been already discounted before intervention. However, the objective was to reassess the opinion of GPs when new information, *i.e.* expert support, was available, regardless of their earlier assessment in the same case.

From a methodological point of view, the use of the volume–time curve as sham information could be questioned. Theoretically, such curves do not show new information to GPs after presentation of the flow–volume curves. Additionally, this is not really ''usual care'', as most GPs in the Netherlands are trained to look at flow–volume curves rather than volume–time curves. Conversely, the volume–time curve is much more intuitive and may have improved

unconscious performance of spirometry interpretation in the control group. Furthermore, providing the expert panel and GPs in the present study with a fixed cut-off value of <0.7, instead of the lower limit of normal for the FEV1/FVC ratio in the standardised cases may have led to an overestimation of diagnosed airflow obstruction [22]. Further discussion about the pros and cons of using a fixed cut-off value *versus* the lower limit of normal for FEV1/FVC [23] is beyond the scope of the present paper.

Finally, a possible reason why no differences could be demonstrated in diagnostic achievements should be sought in the expert support system used. Although the expert support system used in the present study met the criteria of a good system [15], *i.e.* involvement of the present authors by development, integration through the computer, and the displaying of specific recommendations at the right place and time, it was not actually tested in the target group, *i.e.* GPs, before the study. Therefore, it may not optimally comply with the decision-making process of GPs. The information presented by the system to the GP possibly lacked explanation of exactly what the output means. These are known barriers to the adoption of expert support in primary care [24].

Relation to other studies

A recent systematic review [16] demonstrated the following two relevant issues with respect to expert support systems: 1) the effects of diagnostic expert support systems on GPs' performance were low; and 2) trials evaluating diagnostic systems were scarce. Currently, there are no similar expert support studies available with which to directly compare the present results. It is important to realise that, similarly to ECG, spirometry is a highly complex diagnostic tool in the perception of many GPs. Although a recent study evaluating the ECG interpretation skills of GPs and the value of automatic ECG recorded interpretations [25] seemed promising to compare the present study's results with, it lacked the correct design. In the present study, and similarly to the results of the study by JENSEN *et al.* [25], the PPVs were lower than the NPVs. The highest NPVs were found for the cases with the conditions of asthma and absence of respiratory disease. This probably reflects the fact that it is more difficult for a GP to confirm the presence of a disease than to exclude its presence.

The acuity of GPs' interpretation of test results has been evaluated by others. In 1999, EATON *et al.* [5] had already found that 53% of GPs' interpretation of spirometry test results was judged to be correct according to an expert panel. Recently, RAGHUNATH *et al.* [9] found that the agreement in interpretation of spirometry and peak flow results between nurses, GPs and an expert panel was only 20%. The lower agreement in the latter study could probably be explained by the fact that GPs, as well as nurses, *i.e.* less-trained professionals, assessed a common diagnosis. Furthermore, contrary to GPs and nurses, the expert panel did not have detailed clinical history information to assess their final diagnosis on and, due to a design artefact, interpretation of their study results was difficult. Results of the present study concur with the results of EATON *et al.* [5] and show that, generally, GPs have made progress in the interpretation of test results relevant for respiratory diseases in primary care. The current acuity of GPs' interpretation of test results should weaken earlier reported lack of confidence in the ability to interpret the test results [8, 9].

Unanswered questions and future research

Generally, two questions remain to be answered: 1) how can optimal quality spirometry results in primary care be achieved outside of research settings?; and 2) what is the most effective way to give continuous expert support for the interpretation of spirometry test results, given a situation of optimal quality results [14]? Continuous expert support could be provided by means of consultation or feedback from a chest physician or by means of an expert support system. The results of the present study add to current knowledge that computerised spirometry expert support had no detectable benefit over sham information on GPs' diagnostic achievements and decision-making processes when diagnosing chronic respiratory disease. The comparison of support from a chest physician *versus* computerised expert support for spirometry test results calls for further study.

REFERENCES

1　National Institute for Clinical Excellence (NICE). Chronic obstructive pulmonary disease: national clinical guideline for management of chronic obstructive pulmonary disease in adults in primary and secondary care. *Thorax* 2004; 59: Suppl. 1, 1–232.

2　Global Initiative for Chronic Obstructive Lung Disease. www.goldcopd.com. Date last updated: December 2006. Date accessed: November 15, 2006.

3　Caramori G, Bettoncelli G, Tosatto R, *et al.* Underuse of spirometry by general practitioners for the diagnosis of COPD in Italy. *Monaldi Arch Chest Dis* 2005; 63: 6–12.

4　Johns DP, Burton D, Walters JA, Wood-Baker R. National survey of spirometer ownership and usage in general practice in Australia. *Respirology* 2006; 11: 292–298.

5　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: 416–423.

6　Poels PJP, Schermer TRJ, 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: 81–87.

7　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: 391–396.

8　Bolton CE, Ionescu AA, Edwards PH, Faulkner TA, Edwards SM, Shale DJ. Attaining a correct diagnosis of COPD in general practice. *Respir Med* 2005; 99: 493–500.

9　Raghunath AS, Innes A, Norfolk L, *et al.* Difficulties in the interpretation of lung function tests in the diagnosis of asthma and chronic obstructive pulmonary disease. *J Asthma* 2006; 43: 657–660.

10　Chavannes N, Schermer T, Akkermans R, *et al.* Impact of spirometry on GPs' diagnostic differentiation and decision-making. *Respir Med* 2004; 98: 1124–1130.

11　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: 2443–2447.

12　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: 945–952.

13　Schermer TR, Jacobs JE, Chavannes NH, *et al.* Validity of spirometric testing in a general practice population of patients with chronic obstructive pulmonary disease (COPD). *Thorax* 2003; 58: 861–866.

14 Poels PJ, Schermer TR, van Weel C, Calverley PM. Spirometry in chronic obstructive pulmonary disease. *BMJ* 2006; 333: 870–871.

15 Purcell GP. What makes a good clinical decision support system. *BMJ* 2005; 330: 740–741.

16 Garg AX, Adhikari NKJ, McDonald H, *et al.* Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005; 293: 1223–1238.

17 SpidaXpert. Become an Expert in Spirometry. www.spirxpert.com. Date last updated: February 2007. Date last accessed: January 15, 2007.

18 Poels PJP, Schermer TRJ, Akkermans RP, *et al.* General practitioners' needs for ongoing support for the interpretation of spirometry tests. *Eur J Gen Pract* 2007; 13: 16–19.

19 van Weel C, de Grauw W. Family practices registration networks contributed to primary care research. *J Clin Epidemiol* 2006; 59: 779–783.

20 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

21 Knottnerus JA, van Weel C, Muris JWM. Evidence base of clinical diagnosis: evaluation of diagnostic procedures. *BMJ* 2002; 324: 477–480.

22 Hardie JA, Buist AS, Vollmer WM, Ellingsen I, Bakke PS, Mørkve O. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J* 2002; 20: 1117–1122.

23 Schermer TR, Quanjer PH. COPD screening in primary care: who is sick? *Prim Care Respir J* 2007; 16: 49–53.

24 Short D, Frischer M, Bashford J. Barriers to the adoption of computerised decision support systems in general practice consultations: a qualitative study of GPs' perspectives. *Int J Med Inform* 2004; 73: 357–362.

25 Jensen MSA, Thomsen JL, Jensen SE, Lauritzen T, Engberg M. Electrocardiogram interpretation in general practice. *Fam Pract* 2005; 22: 109–113.