ERJ Express. Published on June 27, 2007 as doi: 10.1183/09031936.00012007

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

Patrick JP Poels, Tjard RJ Schermer, Daan PA Schellekens, Reinier P Akkermans, Pieter F de Vries Robbé, Alan Kaplan, Ben JAM Bottema, Chris van Weel

Short Title: Spirometry expert support in general practice

Patrick JP Poels

General Practitioner

Radboud University Nijmegen Medical Centre, Department of General Practice (117), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 33 15

Fax ++31 24 354 18 62, p.j.p.poels@hag.umcn.nl

Tjard RJ Schermer

Senior researcher

Radboud University Nijmegen Medical Centre, Department of General Practice (117), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 33 15

Fax ++31 24 354 18 62, t.schermer@hag.umcn.nl

Daan PA Schellekens

MD

Radboud University Nijmegen Medical Centre, Department of General Practice (117), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 33 15

Fax ++31 24 354 18 62, d.schellekens@hag.umcn.nl

Reinier P Akkermans

Statistician

Radboud University Nijmegen Medical Centre, Department of General Practice (117), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 33 15

Fax ++31 24 354 18 62, r.akkermans@ives.umcn.nl

Pieter F de Vries Robbé

Professor of Medical Informatics

Radboud University Nijmegen Medical Centre, Department of Medical Informatics (152), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 32 51

Fax ++31 24 361 3504, p.devriesrobbe@mi.umcn.nl

Alan Kaplan

General Practitioner

17 Bedford Park Avenue, Richmond Hill , ON L4C 2N9, Ontario, Canada

Phone: 905-883-1100, Fax: 905-884-1195, FOR4KIDS@sympatico.ca

Ben JAM Bottema

General Practitioner

Radboud University Nijmegen Medical Centre, Department Postgraduate Training (166), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone ++31 24 361 53 00

Fax++31 024 361 95 53 b.bottema@voha.umcn.nl

Chris van Weel

Professor of General Practice

Radboud University Nijmegen Medical Centre, Department of General Practice (117), PO box 9101, 6500 HB, Nijmegen, The Netherlands, Phone +31 24 361 33 15

Fax ++31 24 354 18 62, c.vanweel@hag.umcn.nl

Correspondence to: p.j.p.poels@hag.umcn.nl

Word count: 3368

Number of tables: 4 (plus 1 online depository table)

Number of figures: 2

Number of references: 25

Abstract

This 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.

We performed a cluster-randomised controlled trial in 78 GPs who completed 10 standardised paper case descriptions each. 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 were 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 non-medication changes. Effects were expressed as odds ratios (OR) with 95% confidence intervals.

There were no differences between the expert support and the control group in the agreement between GP and expert panel diagnosis of COPD (OR=1.08 [95% CI 0.70-1.66]), asthma (OR=1.13 [95% CI 0.70-1.80]), and absence of respiratory disease (OR=1.32 [95% CI 0.61-2.86]). A higher rate of additional diagnostic tests was observed in the expert support group (OR= 2.5 [95% CI 1.17-5.35]).

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

Trial Number: http://www.clinicaltrials.gov/ct/show/NCT00131157?order=1

Keywords: Diagnosis Computer-Assisted, Expert Systems, Family Practice, Spirometry

Introduction

Although all major 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 use spirometry consequently in 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] lack of time and practice support to fit spirometry into the daily practice routine,[6] absence of a spirometer in the practice,[7] and GP's lack of confidence in the ability to interpret the test results.[8;9] A recent survey showed that a third of Australian GPs interpreted less than one spirometry test per week.[4] Due to this low prevalence of test interpretations, it seems difficult for GPs to become an expert in this.

We have previously demonstrated the influence of spirometry on GPs' diagnostic achievements and management decisions in a non-randomised simulation study.[10] Other recent non-randomised studies confirm that spirometry increases diagnostic rates of chronic respiratory disease and may lead to management changes in a general practice population.[11;12] 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 we observed that the most relevant indices as measured by trained general practice staff were comparable with those measured in pulmonary function laboratories.[13]

Therefore, once GPs have had initial spirometry training and spirometry equipment and test validity is adequate, the next step to improve implementation of spirometry in general practice is to arrange for a possibility to receive continuous advice and support for the test interpretation.[14] This could be done 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 kind 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 the diagnostic achievements of GPs, and on GPs' decision-making when diagnosing chronic respiratory disease.

Methods

Study design

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

Participants

GPs from the catchment's area of our academic hospital and from a specific general practice network of our department[19] were invited by postal mailing to participate.

Intervention

GPs were randomly allocated to the (i) computerised spirometry expert interpretation support group, or (ii) the control group. GPs in the expert support group received the spirometry test results, flow-volume curve, plus the graphical interpretation and the textual interpretative notes. GPs in the control group received the spirometry test results, flow-volume curve, plus the volume-time curve (figure 1).

The spirometry expert system (SpidaXpert®, Micro Medical Ltd, Rochester, UK),[17] contains a diagnostic algorithm based on pre- and postbronchodilator FEV_1/FVC and FEV_1 values and the accompanying age, sex, and origin specific predicted values. The expert interpretation module in SpidaXpert® had been developed with funding of the Dutch Asthma Foundation by a group of independent experts (http://www.spirxpert.com/spirxpertgroup.htm). 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, that 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. We introduced sham information in the control group to be able to compare GPs re-assessment of a diagnosis in the control group in the same way as in the expert support group. Sham information is in fact a "placebo-effect" as we presented no new data to these GPs; we presented earlier data (i.e. the flow-volume curve) in each case again in another way (i.e. the volume-time curve). Although clearly important to evaluate the quality of forced expiratory manoeuvres (i.e., end of test criteria),[20] from a diagnostic point of view the volume-time curve does not add relevant new information to the information provided by the flow-volume curve and the numeric test results. Prior to the study, participants were informed that they would receive additional information on the spirometry and asked to reconsider their diagnosis. No further specification was given of the nature or the background of that information.

INSERT FIGURE 1 HERE

Standardised case descriptions and gold standard

Based on our experiences in a previous study,[10] we knew beforehand that GPs are quite able to diagnose common respiratory disease patterns, whereas rare pathology 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) growths with age. That is the reason for including case descriptions of adult patients only, with a special focus on the 50-60 year age category. This category reflects daily practice patterns in primary care. The case descriptions - in which a GP would use spirometry as a diagnostic test – were: COPD (classified as GOLD stage 1 (n=1); GOLD stage 2 (n=1); GOLD stage 3 (n=2))[2]; asthma (n=2); allergic asthma (n=1); lung fibrosis

(n=1); no respiratory disease (n=1); incorrect test manoeuvre (n=1); exercise-induced asthma (n=1) [example case] (see online depository).

At inclusion, a research assistant visited the participating GPs in their practice. During a 90 minutes audio taped 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 re-consider their diagnosis and management after the intervention. An example of the case structure is depicted in Figure 2. Due to time limitation we asked only for specific medication and non-medication changes after the intervention in cases with already diagnosed respiratory disease (6 out of 10 cases).

INSERT FIGURE 2 HERE.

Before their use in the study, the cases were judged by an expert panel consisting of two chest physicians, a GP (PP) with specific expertise in spirometry and a health scientist (TS). The panel consensus diagnoses served as “the gold standard” in the subsequent evaluation of GPs’ diagnostic achievements.

The whole approach was piloted in 4 GPs. Shortly after the first six visits, we adjusted the case set 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 6 GPs (equally divided over the 2 groups).

Primary and secondary outcome measures

Difference in proportion of agreement of the cases' diagnosis between GP's and the expert panel judgement before and after interpretation of spirometry served as the primary outcome and were directed at 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 GP's decision-making process: (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) GP's certainty of the working diagnosis (self-scored between 0-10, [0 = uncertain, 10=certain]); (4) GP's perception of severity of the working diagnosis (self-scored between 0-10, [0=no severe disease, 10=severe disease]); (5) probability of referral to secondary care (yes/no); (6) probability of medication and non-medication changes. Medication change included either to stop or lower of inhaled corticosteroids or bronchodilators or to start of bronchodilators, inhaled corticosteroids, oral corticosteroids, and combination drugs. Non-medication included giving smoking cessation advice.

Sample size

Calculation of the sample size was based on an estimated relevant proportion 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], α=0.05, 1-β=0.80, and an intra-cluster correlation r=0.18, 31 GPs were required in each randomisation group. To allow for drop-outs and subgroup analyses we aimed at including a minimum of 70 GPs.

Randomisation

The research assistant used restricted randomisation (minimisation) with a computer program on a laptop computer using three stratification factors: GP's prior experience with the specific computerised spirometry interpretation support package (yes/no), the average number of spirometry tests a GP reported to interpret per week, and GP's experience (years) with spirometry. The researchers and the statistician (RA) were blinded while assessing and reporting all outcomes.

Statistical analysis

Agreement between GP's and the expert panel judgement was expressed as percentages with 95% confidence intervals (95%CI). Multilevel regression logistic modelling was used to account for the intra-cluster 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. We performed multilevel logic analyses for dichotomous variables and multilevel regression analyses for continuous variables in SAS V8.2 for Windows (SAS Institute Inc, Cary USA 1999-2001). Odds ratios (OR) with 95%CI 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 (further referred to as PPV and NPV, respectively), and the diagnostic odds ratio (DOR)[21] with 95% confidence intervals were calculated for GP judgements of COPD, asthma, rare respiratory pathology and no respiratory disease after the intervention. Odds ratios with 95% confidence intervals were also used to evaluate differences in indicators GPs' decision-making process.

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

Results

Baseline characteristics of GPs

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

Table 1 Baseline characteristics of all randomised GPs.

	Expert support group (n=36 GPs)	Control group (n=42 GPs)
Type of practice, n (%)		
-single handed	2 (5)	4 (10)
-duo	9 (25)	9 (21)
-group (≥ 3 GPs)	15 (42)	19 (45)
-multidisciplinary health care centre	10 (28)	10 (24)
Gender, % male	64	57
GP's experience with spirometry in years, *mean (SD)*	5.5 (4.3)	5.3 (3.3)
No. of spirometry results interpreted per week, *mean (SD)*	1.4 (0.8)	1.4 (0.7)
Prior experience with expert support, % yes	47	36

Primary outcome: diagnostic achievements by GPs

GPs assessed in total 774 cases; 357 cases in the expert support group and 417 cases in 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 before intervention 66.0% (expert support) versus 65.9% (control) and after intervention 68.5% (expert support) versus 63.5% (control).

Table 2 Agreement on case diagnoses between GPs' and expert panel judgement before and after intervention.
The Odds Ratio expresses the difference in GP's judgement before and after intervention, expert support compared to control.

	Expert support group (n=357 cases)		Control group (n=417 cases)		Expert panel	Odds ratio (95% CI)
	intervention		intervention		judgment	
GP judgement presence of:	before	after	before	after		
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)

NA = not available, 95%-CI = 95% confidence interval

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

Table 3 Sensitivity, specificity, predictive values, and diagnostic odds ratios for GPs' judgment 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
Positive predictive value,%	70.7	78.9	33.3	61.9
Negative predictive value,%	85.5	92.6	90.1	94.0
Diagnostic odds ratio	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
Positive predictive value,%	71.1	78.0	0.0	60.9
Negative predictive value,%	83.1	89.8	89.9	93.7
Diagnostic odds ratio	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

NA = not available, 95%-CI = 95% confidence interval

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 2.5 [95%CI 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 non-pulmonary drugs) between the study groups.

Table 4 Impact of the intervention on six indicators of GPs' decision-making process.
*The Odds Ratio expresses the difference in an indicator of GPs' decision-making process before and after the intervention, expert support compared to control. Statistically significant differences (p<0.05) are printed in **bold**.*

	Expert support group (N=357)		Control group (N=417)		Odds Ratio	95% CI
	intervention		intervention			
Indicators	before	after	before	after		
(1) Additional diagnostic tests,*%*	70.2	5.2	75.3	3.0	**2.5**	(**1.17 – 5.35)**
- X ray 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)
- prednisolon 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[a]	1.1	0.0	0.7	0.0	NA	NA
(2) Width of differential diagnoses, *mean (SD)*	2.2 (1.0)	1.7 (0.9)	2.3 (1.0)	1.7 (1.0)	1.0	(0.81 – 1.19)
(3) Certainty of diagnosis, *mean (SD)*	6.8 (2.0)	7.1 (1.9)	7.3 (1.9)	7.3 (1.8)	1.0	(0.57 – 1.43)
(4) Perception of severity of diagnosis, *mean (SD)*	6.0 (2.2)	5.9 (2.3)	6.3 (2.1)	6.3 (2.1)	1.0	(0.57 – 1.42)
(5) Referral rate,*%*	18.6	1.7	17.8	2.5	1.0	(0.60 – 1.66)
(6) Medication & non-medication changes[b],*% yes*	-	76.2	-	69.1	1.44	(0.80 – 2.59)
- 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)
- Non-medication changes						
- smoking cessation advice	NA	30.1	NA	33.7	0.85	(0.57 – 1.26)

NA = not available, 95%-CI = 95% confidence interval
a includes urine test, gastroscopy, ergometry, blood pressure, temperature and oxygen saturation
b this information was available for 6 out of 10 cases (expert support n=216, control group n=252 cases)

Subgroup analyses

Neither GP's experience with spirometry (OR 1.02 [95%CI 0.97 - 1.06]), nor GP's prior experience with expert support (OR 0.97 [95%CI 0.72 - 1.31]), nor GP's number of interpreted spirometry test per week (OR 1.02 [95%CI 0.84 - 1.23]) were associated with the effectiveness of expert support, as their agreements with the expert panel was not different before intervention. If GPs interpreted more spirometry test per week and had prior experience with expert support the probability of agreement with the expert panel before intervention increased. It 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 process.

Strengths of the study

This is the first diagnostic study that assesses the impact of a commercially available computerised expert support for spirometry in a randomised simulation study in primary care. The study used standardized patients, which means that all participants were faced with the same diagnostic challenges. This can only be achieved in an in-vitro design as in real life practice patient mix would cause difficult to capture variation in diagnostic challenges.

The standardised complex and original method that we used to assess this impact had been used before in a non-randomised design.[10] Based on that information we were able to create a balanced mixture of cases relevant for GPs. We focussed stronger on the confirmative role of spirometry 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 (RA).

Subgroup analyses did not show any difference in baseline diagnostic achievements of GPs on prior experience with spirometry, on GP's prior experience with this expert support system, and on the number of spirometry test 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

Our trial has some limitations. In a diagnostic assessment of chronic respiratory disease, GPs' consideration to perform spirometry in case of a intermediate prior probability of disease is a great diagnostic step.[21] This step was already foreseen in our study design. The next step of diagnostic refinement does not seem to influence the posterior probability extensively. In our 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 our assumptions in the power calculation (50% correct diagnoses without expert support). Probably instruction and support for these GPs had not been effective, as these GPs could already be considered as "experts" due to prior participation in other studies or postgraduate spirometry training programs from our department. Therefore, the expert system had hardly additional value and could be considered as a sort of luxury appendix for these GPs.

We found a large within group difference for ordering additional diagnostic tests, which may be an effect of the study design: because GPs expected from us the results of their diagnostics already discounted before intervention they hardly reassessed their diagnostics after intervention. However, the objective was to re-assess GP's opinion when new information (i.e. expert support) was available, regardless of their earlier assessment in the same case.

From a methodological point of view, one could question the use of the volume-time curve as sham information. Theoretically these curves do not add new information to GPs after presentation of the flow-volume curves. On the one hand this is not really 'usual care' as most GPs in our country are trained to look at flow-volume curves rather than volume-time curves. On the other hand, 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 our study with the fixed cut-off <0.7 in stead of the lower limit of normal (LLN) for the FEV_1/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 the fixed cut-off *versus* the LLN for FEV_1/FVC[23] is beyond the scope of this paper.

Finally, a possible reason why we could not demonstrate differences in diagnostic achievements should be sought in the used expert support system. Although the expert support system we used meets the criteria of a good system[15] – involvement of authors by development, integrated in computer, displaying 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 what the output means exactly. These are known barriers to the adoption of expert support in primary care.[24]

Relation to other studies

A recent systematic review demonstrated two relevant issues with respect to expert support systems: (1) the effects of diagnostic expert support systems on practitioners performance was low, and (2) trials evaluating diagnostic systems were scarce.[16] Currently, there are no similar expert support studies available to compare our results with directly. It is important to realise that - like electrocardiography (ECG) - spirometry is a highly complex diagnostic tool in the perception of many GPs. Although a promising idea to evaluate the ECG interpretations skills of GPs and the value of automatic ECG recorded interpretations, a recent study lacked the right design to compare our results with.[25] Like the results from the study from Jensen et al.[25] we found in our study that the positive predictive values were lower than the negative predictive values. For the cases with the conditions asthma and

absence of respiratory disease we found the highest NPVs. This probably reflects the fact 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 had been evaluated by others. In 1999, Eaton et al. found already that 53% of GPs' interpretation of spirometry test results was judged correct according to an expert panel.[5] Recently, Raghununath et al. found that the agreement in interpretation of spirometry and peak flow results between nurses, GPs and an expert panel was only 20%.[9] The lower agreement in this 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 the nurses and GPs, the expert panel did not have detailed clinical history information to assess their final diagnosis on. Due to design artefact, interpretation of their study results is difficult. Results of our study confess Eaton's[5] results 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 GP's interpretation of the 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, the first question to answer is how to achieve optimal quality spirometry results in primary care outside of research settings. The second question to answer is what is the 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 of feedback from a chest physician or by means of an expert support system. Results of the current study add to knowledge that computerised spirometry expert support had no detectable benefit over sham information on GPs' diagnostic achievements and decision-making process when diagnosing chronic respiratory disease. Comparison of support from a chest physician versus computerised expert support for spirometry test results calls for further studies.

Acknowledgements

We thank all participating GPs; we thank Yvonne Heijdra (PhD) and Johan Molema (PhD), chest physicians, for their role in the expert panel; Margreet van den Bogart-Jansen (MSc), Antoinette Marks (RN) who helped with the interviews of the GPs; Joke Grootens-Stekelenburg for data management.

Executive steering committee: Henk van den Hoogen, MSc; Annelies Jacobs, MSc PhD; Philip Quanjer, MD PhD; Bart Thoonen, MD PhD.

Funding: This study was supported by grant 3.4.02.18 of the Netherlands Asthma Foundation and grant 920.03.265 of the Netherlands Organisation for Health Research and Development. The authors' work was independent of the funders (the funding source had no involvement).

Ethical approval: This study was approved by the medical ethics review board of the Radboud University Nijmegen Medical Centre.

Competing interests: All authors declare have nothing to declare.

Legends

Figure 1 Participants and intervention

Figure 2 Schematic representation of a case structure

Reference List

(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 I):i1-i232.

(2) Global Initiative for Chronic Obstructive Lung Disease; www.goldcopd.com. Date accessed: November 15 2006.

(3) Caramori G, Bettoncelli G, Tosatto R, Arpinelli F, Visona G, Invernizzi G et al. Underuse of spirometry by general practitioners for the diagnosis of COPD in Italy. Monaldi Arch Chest Dis 2005; 63(1):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(3):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(2):416-423.

(6) Poels PJ, Schermer TR, Jacobs A, Akkermans RP, Hartman J, Bottema BJ et al. Variation in spirometry utilisation between trained general practitioners in practices equipped with a spirometer. Scan J Prim Health Care 2006; 24(2):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(5):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(4):493-500.

(9) Raghunath AS, Innes A, Norfolk L, Hannant M, Greene T, Greenstone M et al. Difficulties in the interpretation of lung function tests in the diagnosis of asthma and chronic obstructive pulmonary disease. J Asthma 2006; 43(9):657-660.

(10) Chavannes N, Schermer T, Akkermans R, Jacobs JE, van de GG, Bollen R et al. Impact of spirometry on GPs' diagnostic differentiation and decision-making. Respir Med 2004; 98(11):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(4):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(5):945-952.

(13) Schermer TR, Jacobs JE, Chavannes NH, Hartman J, Folgering HT, Bottema BJ et al. Validity of spirometric testing in a general practice population of patients with chronic obstructive pulmonary disease (COPD). Thorax 2003; 58(10):861-866.

(14) Poels PJ, Schermer TR, van Weel C, Calverley PM. Spirometry in chronic obstructive pulmonary disease. BMJ 2006; 333(7574):870-871.

(15) Purcell GP. What makes a good clinical decision support system. BMJ 2005; 330(7494):740-741.

(16) Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J et al. Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review. JAMA 2005; 293(10):1223-1238.

(17) SpidaXpert, software for interpreting pulmonary function. www.spirxpert.com. Date accessed: January 15 2007.

(18) Poels PJP, Schermer TRJ, Akkermans RP, Jacobs A, van den Bogart-jansen M, Bottema BJAM et al. General practitioners' needs for ongoing support for the interpretation of spirometry tests. Eur J Gen Pract 2007; 2007(13):16-19.

(19) van Weel C, de Grauw WJC. Family practices registration networks contributed to primary care research. J Clin Epidemiol 2006; 59(8):779-783.

(20) Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A et al. Standardisation of spirometry. Eur Respir J 2005; 26(2):319-338.

(21) Knottnerus JA, van Weel C, Muris JW. Evidence base of clinical diagnosis: Evaluation of diagnostic procedures. BMJ 2002; 324:477-480.

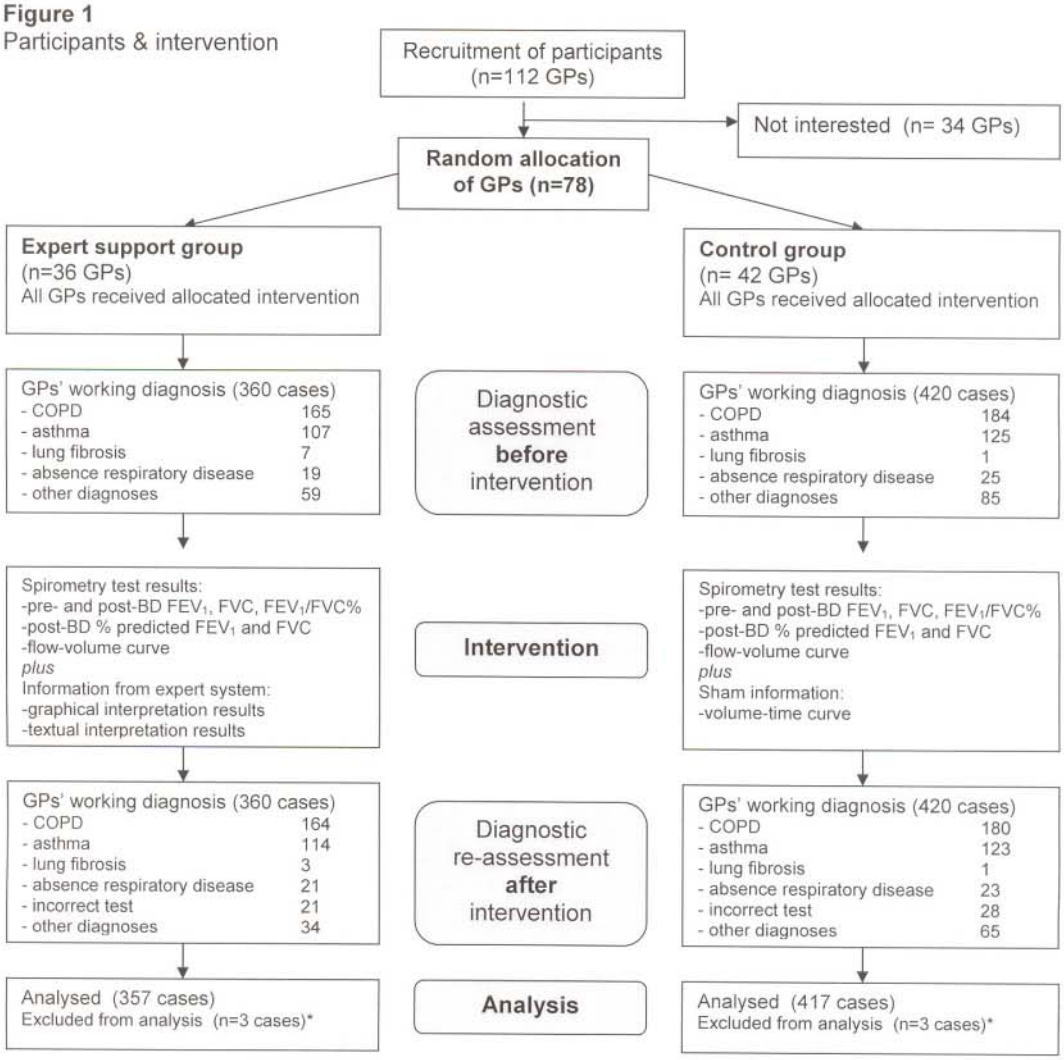
(22) Hardie JA, Buist AS, Vollmer WM, Ellingsen I, Bakke PS, Morkve O. Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. Eur Respir J 2002; 20(5):1117-1122.

(23) Schermer TR, Quanjer PH. COPD screening in primary care: who is sick? Prim Care Respir J 2007; 16(1):49-53.

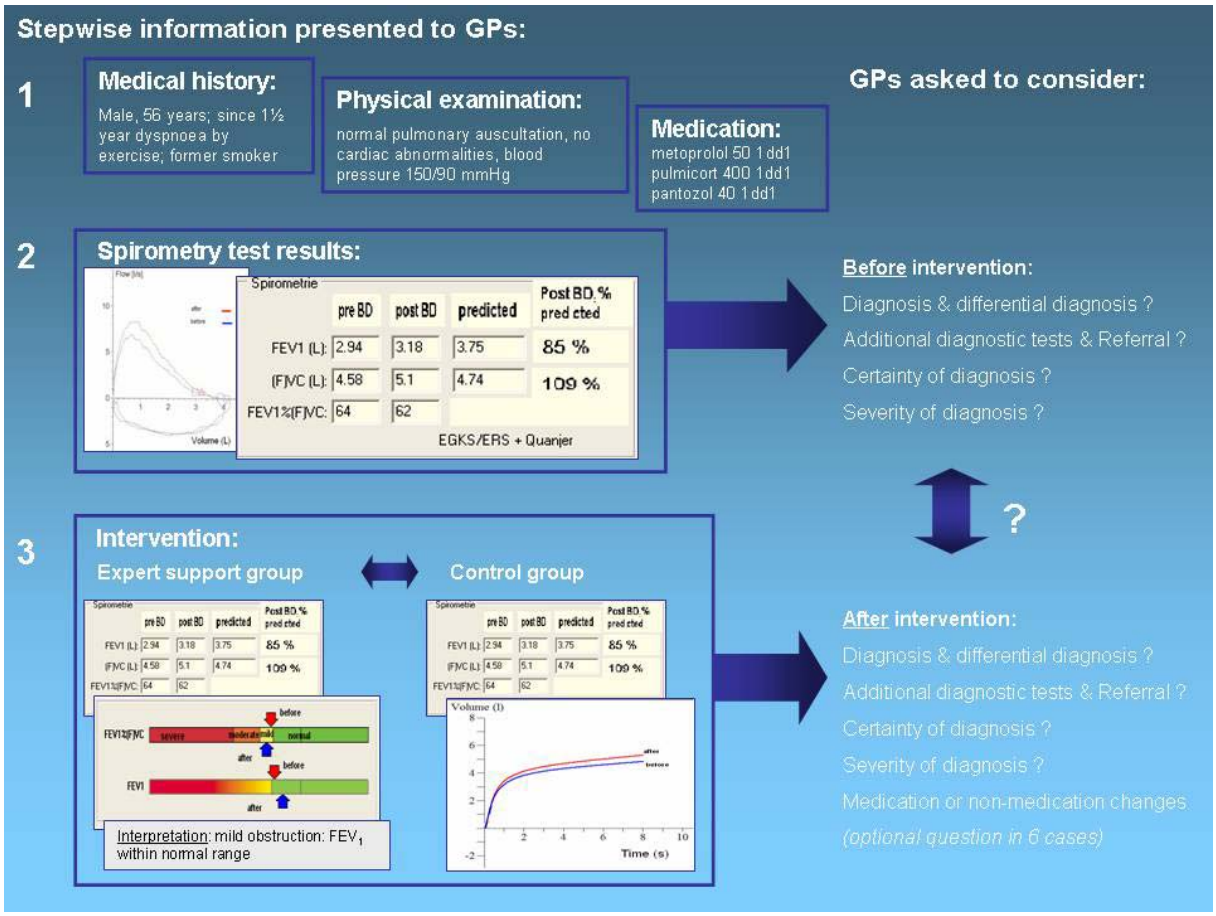
(24) Short D, Frischer M, Bashford J. Barriers to the adoption of computerised decision support systems in general practice consultations: a qualitative study of GPs' perspectives. Int J Med Inform 2004; 73(4):357-362.

(25) Jensen MSA, Thomsen JL, Jensen SE, Lauritzen T, Engberg M. Electrocardiogram interpretation in general practice. Fam Pract 2005; 22(1):109-113.

Figure 1
Participants & intervention



*the first 6 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 consecutive other 72 GPs we switched the case set between these 2 cases: the case of "absence of respiratory disease" was included for them in the final case set. Therefore, we did not have information 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=3 GPs).
Used abbreviations:
GP = general practitioner, COPD = chronic obstructive pulmonary disease, BD = bronchodilatation, FEV_1 = forced expiratory volume in first second, FVC = forced vital capacity.


Stepwise information presented to GPs:
1
Medical history:
Male, 56 years; since 1½ year dyspnoea by exercise; former smoker
Physical examination:
normal pulmonary auscultation, no cardiac abnormalities, blood pressure 150/90 mmHg
Medication:
metoprolol 50 1 dd1
pulmicort 400 1 dd1
pantozol 40 1 dd1
GPs asked to consider:
2
Spirometry test results:
Spirometrie
pre BD
post BD
predicted
Post BD %pred cted
FEV1 (L): 2.94 3.18 3.75 85 %
(F)VC (L): 4.58 5.1 4.74 109 %
FEV1%(F)VC: 64 62
EGKS/ERS + Quanjer
Before intervention:
Diagnosis & differential diagnosis ?
Additional diagnostic tests & Referral ?
Certainty of diagnosis ?
Severity of diagnosis ?
?
3
Intervention:
Expert support group
Control group
Interpretation: mild obstruction: FEV1 within normal range
Volume (l)
Time (s)
After intervention:
Diagnosis & differential diagnosis ?
Additional diagnostic tests & Referral ?
Certainty of diagnosis ?
Severity of diagnosis ?
Medication or non-medication changes
(optional question in 6 cases)