Spirometry in patients with COPD:

Focus on expiratory ànd inspiratory parameters

Frank Visser

ISBN: 978-94-6191-153-7

Printed by: Ipskamp Drukkers, Enschede
Cover: Eline Visser
Layout: Theo Hafmans Fotografie, Berg en Dal

Spirometry in patients with COPD:

Focus on expiratory ànd inspiratory parameters

Een wetenschappelijke proeve op het gebied van de Medische Wetenschappen

Proefschrift

Ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof.mr. S.C.J.J. Kortmann
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 10 februari 2012
om 10.30 uur precies

door

Frans JohanVisser

geboren op 26 november 1950
te Terneuzen

Promotor:

Prof.dr. P.N.R. Dekhuijzen

Copromotor:

Dr. Y.F. Heijdra

Manuscriptcommissie:

Prof.dr. M.T.E. Hopman

Dr. J. Deinum

Prof.dr. P. Sterk, Amsterdam Medisch Centrum

Paranimfen:

Jos Schulte

Fleur Visser

Contents

1

General introduction

General introduction

Definitions of COPD, asthma and airway obstruction.
This paragraph discusses the definitions of Chronic obstructive pulmonary disease (COPD), asthma and airway obstruction. First of all, according to the World Health Organization, COPD is defined as: "Chronic obstructive pulmonary disease (COPD) is a lung disease characterized by chronic obstruction of lung airflow that interferes with normal breathing and is not fully reversible. The more familiar terms 'chronic bronchitis' and 'emphysema' are no longer used, but are now included within the COPD diagnosis. COPD is not simply a "smoker's cough" but an under-diagnosed, life-threatening lung disease. [1]". This definition differs substantially from the definition of the American Thoracic Society and European Respiratory Society (ATS-ERS) which resembles the definition of the Global Initiative for Obstructive Lung Disease (GOLD): "Chronic obstructive pulmonary disease (COPD) is a preventable and treatable disease characterized by airflow limitation that is not fully reversible. The airflow limitation is usually progressive and is associated with an abnormal inflammatory response of the lungs to noxious particles or gases, primarily caused by cigarette smoking. Although COPD affects the lungs, it also produces significant systemic consequences."[2,3]

Secondly, the Global Initiative for Asthma (GINA) defines asthma by its clinical, physiological and pathological characteristics. The main physiological feature of asthma is episodic airway obstruction characterized by expiratory airflow limitation [4].

Finally the definition of airway obstruction is disputed by the ERS-ATS and the GOLD. The ATS-ERS states that obstruction is present when the forced expiratory volume in the first second (FEV1) and FEV1/VC (Vital Capacity) are both smaller than the lower limits of normal (LLN), being 1.64 * SD of the Residual Standard Deviation [5]. The GOLD guideline postulates that a post bronchodilator FEV1/FVC (Forced Vital Capacity) smaller than 0.7 confirms the presence of airflow limitation that is not fully reversible [3].
The fact that the GOLD definition of obstruction with fixed limits may lead to a bias in the diagnosis of COPD has already been recognized [5,7,8]. In

this thesis we will use the ERS-ATS definition of obstruction as postulated by Pellegrino et al [5]. It is however surprising and confusing that in another document of the ATS-ERS, the fixed ratio of FEV1/FVC of 0,7 is adopted from the GOLD guidelines [3]. Nevertheless, based on these definitions, pulmonary function tests are mandatory for making the diagnosis on the obstructive lung diseases Asthma and COPD.

Discussion on severity of COPD

Several classifications have been developed according to the severity of COPD. The GOLD guideline, for instance, uses a simple classification based on FEV1 as the percentage of the predicted value. This classification is shown in Table 1.

Table 1

COPD severity	GOLD stage	Definition FEV1 as percentage predicted value in case of obstructive lung disease
Mild	I	FEV1 ≥ 80%
Moderate	II	50≥ FEV1 < 80%
Severe	III	30≥FEV1<50%
Very severe	IV	FEV1<30% **or** FEV1<50% plus respiratory failure

However, this classification, as well as the ATS-ERS spirometric classification [3], is not a very good predictor for prognosis of life expectancy nor for the disease burden in a single patient [9,10], since additional factors, which may influence lung function, are not taken into account. These factors include other lung function parameters, transfer capacity of carbon monoxide, exercise tests, health status, dyspnea measurements, exacerbation frequency and the level of emphysema measured by CT-scanning and biomarkers (in exhaled air or condensate of exhaled air and blood). In order to make a better prediction of life expectancy, the bode index is introduced [9,11]. Bode is an acronym for Body mass index, Obstruction, Dyspnea and Exercise measured by the six minutes walking test. Nevertheless, this instrument is not able to accurately predict life expectancy nor the end of a life phase of a subject with severe COPD [12]. In this thesis we use the GOLD definition on severity of COPD because of its simplicity.

Expiratory versus inspiratory parameters

FEV1 is a very important and widely used expiratory lung function parameter. It is used as a measure of severity of airflow obstruction, to test airway hyperresponsiveness and to test airway reversibility. It is, however, disappointing that this expiratory parameter does not correlate well with the dyspnea feeling of the patient. Dyspnea is one of the major complaints in subjects with COPD. The weak association between the change in dyspnea and the change in FEV1 or FVC after the use of bronchodilators [13-16] may be due to

1. Airway compression during forced expiration
2. The absence of dyspnea in rest in patient with stable COPD, and therefore an inability to improve after the use of bronchodilators.

Inspiratory parameters like Forced Inspiraty Volume in the first second (FIV1), however, are not influenced by the above mentioned airway compression. Therefore we are interested in the value of inspiratory parameters in the management of COPD. Inspiratory parameters could be more sensitive to interventions like bronchodilators because the airway collapse during forced expiration may obscure benefits of the interventions. In the literature, O'Donnel and Taube [16,17]give support for the idea that inspiratory parameters may be more sensitive to find significant improvements after bronchodilators in patients with COPD. Inspiratory parameters can be obtained after a fast expiration (Fe) or after a slow expiration (Se). However, it is not clear yet which of these two methods would be best in order to measure inpiratory parameters. Therefore we compared these methods and presented the results in chapter two.

Reversibility and random variation

Another important item to discuss is the reversibility or responsiveness (we use these terms interchangeably in this thesis) of lung function parameters after an intervention like bronchodilation or pursed-lips breathing (PLB). There is no consensus about the way to express reversibility, not even for the most known parameter i.e. FEV1 [5]. On the other hand there is consensus that reversibility within one subject must exceed the random variation= natural variation for that parameter [5]. The random or natural variation can be represented in a so called Bland-Altman plot[18]. To understand these plots we present in Figure 1 a simple example and Figure 2 an illustration of two different types of scatter.

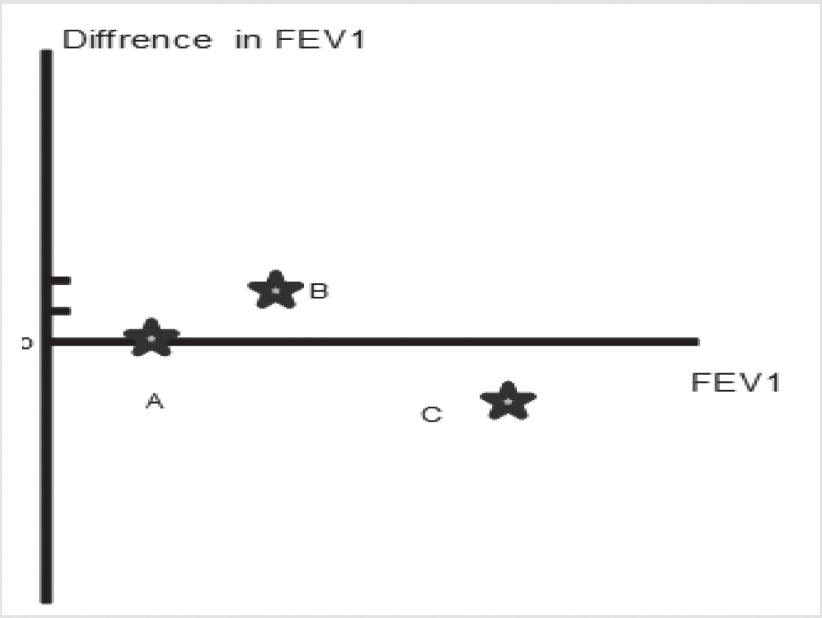


***Figure 1**. A simple example of a Bland Altman plot.*

Patient A, a man with severe COPD was tested two times with a time interval of one hour, he had on both occasions an FEV1 of 0,5 L. (horizontal axis 0,5 L and vertical axes no difference=0. Patient B a women had a FEV1 0f 1 L the first time and 1.2 L an hour later (horizontal axis FEV1 1,1 L being the average of 1 and 1,2 L, snd + 0.2 on the vertical axis). Patient, C a man with moderate COPD expires the first time a FEV1 of 2 L one hour later 1,8 L (difference -0,2 L on the vertical axis and 1,9 L on the horizontal axis). In real a study has to be done with some hundreds of test-retest measurements in order to visualize the scatter for the natural variation in subjects with COPD. If the mean scatter is the same on all values of the FEV1 (horizontal axes), this is called a homoscedastic scatter. But it is also possible that f.e. the mean scatter is increasing when the FEV1 value increases, this is called a heteroscedastic scatter. The latter is actually unwanted because for each patient with another FEV1 there is also another magnitude for the natural scatter.

Figure 2 is an illustration of these two different types of scatter.

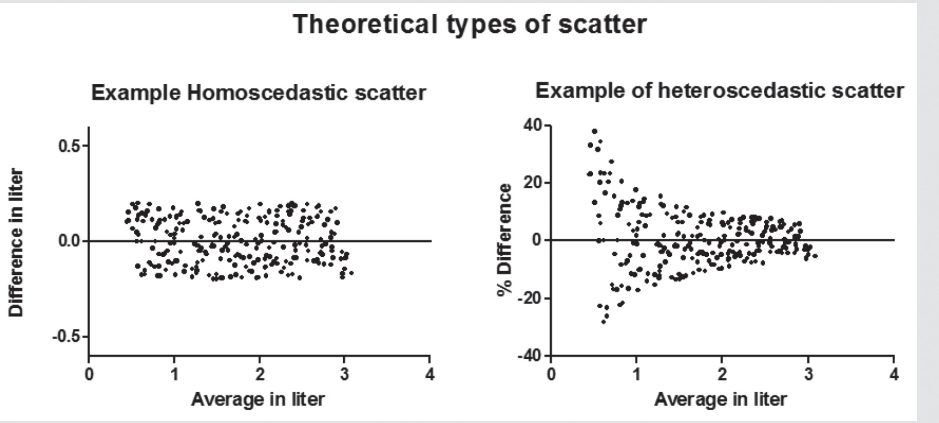


Figure 2. *Bland and Altman plots with theoretical types of scatter.*

The vertical axis shows the test (T) retest (R) difference (R-T); the horizontal axis shows the average (T+R/2). Each point corresponds with one test-retest pair. Left panel: dataset with a random scatter not dependent on the parameter value. Right panel: same dataset, but now the percent change from the average parameter value is shown; the scatter is highly dependent on the parameter value. In the left panel, the scatter can be described with just one value, e.g., the coefficient of repeatability (CR). In the right panel, the scatter cannot be precisely described with one value because the scatter on the low and high parameter values will not precisely reflect a fixed CR. Thus, for describing random variation, a homoscedastic scatter is preferred.

Reversibility can be expressed in different ways

- Absolute difference i.e. in liters, the ATS-ERS taskforce choose for at least 200 ml improvement in FEV1 as a part of their reversibility criterion but they acknowledge that there is no consensus about this [5] .
- Percentage of predicted value
- Percentage of initial value, the ATS-ERS taskforce recommend also at least 12% increase with respect to initial value [5].

The magnitude of the scatter of random variation can also depend on the initial parameter value. Therefore, not only the magnitude but also the type (homoscedastic or heteroscedastic) scatter has to be known. In the literature, recommendations on the method to be used to detect the type of scatter of a parameter have been published [5,18], but to our knowledge

no literature is available that implements this method for lung function parameters. For the magnitude of the scatter there is an advice to use the coefficient of repeatability instead of the more commonly used coefficient of variation [18]. The implementation of this method for detecting the type and magnitude of the scatter for FEV1 is described in chapter three and in chapter four this is done for different inspiratory lung function parameters.

Treatment goals for COPD

In general, treatment goals for patients with COPD aim to improve the prognosis and the quality of life. To improve the quality of life and consequently to minimize the burden of disease, several treatment options are available. Besides the most important treatment options, including giving up smoking, exercise training and maximal bronchodilatation, pursed lips breathing (PLB) is one of the treatment options.

ThePursed lips breathing technique in which the patients expires quietly with nearly closed lips is used as an item of patient education in rehabilitation programs [19,20]. PLB may improve pulmonary gas exchange [21,22], reduces the breathing frequency (BF) and decreases hyperinflation [23,24]. A decrease in dyspnea and an increase in tidal volume are other consequences of PLB in patients with moderate-to-severe COPD [20]. Additionally, a faster recovery from dyspnea and a slower respiratory rate were found after walking with PLB [25]. Because of the importance of inspiratory parameters in this thesis, we studied the effects of PLB on these parameters. In chapter five we describe the effect of PLB in patients with severe-to-very-severe COPD (GOLD stages III and IV) on inspiratory parameters as well as FEV1, forced vital capacity (FVC), oxygen saturation, end-tidal CO2 tension (ET-CO2), BF, and dyspnea. In chapter six we discuss our results of short acting bronchodilator use on inspiratory parameters.

Optimal lung function testing.

Finding the optimal diagnostic work-up in patients with obstructive lung disease is challenging and will be discussed in chapter six. A physician who routinely orders generously PFTs in the work-up of patients runs the risk of unnecessary testing. However, a physician who orders tests more sparingly runs the risk of unnecessary outpatient visits. In view of the high incidence of patients with obstructive lung diseases, it is important to find the optimal diagnostic work-up in each of these patients. To this end, we have developed

a diagnostic protocol that can be jointly used by physicians and pulmonary function assistants. In chapter 6 we examined whether following a PFT protocol is more cost effective than the classical way of test ordering.

Aim and outline of this thesis.

In **chapter two** we aim to formulate recommendations for the measurement of the FIV1 manoeuvre in COPD patients by focusing on the following items:
(1) The difference in FIV1 when performed after a slow (FIV1-Se) or a fast (FEV1-Fe) expiration and its relation to the severity of COPD.
(2) The correlation between, and interchangeability of, FIV1-Se and FIV1-Fe.
(3) The variability of FIV1-Se and of FIV1-Fe and its dependence on the severity of COPD.
(4) The optimal number of FIV1-Se manoeuvres in order to get an acceptable value for the FIV1, and whether this number is dependent on the severity of COPD.

In **chapter three** we will investigate which type and magnitude of scatter applies to the (absolute and percent) changes in FEV1. Thereafter, we will hold these result against the recommendation of the ERS-ATS taskforce for interpretative strategy of PFT's

In **chapter four** we will investigate which type of scatter applies to the (absolute and percent) changes in inspiratory parameters (FIV1, Inspiratory Capacity (IC), Peak Inspiratory Flow (PIF) and Maximal Inspiratory Flow at 50% (MIF50)). Next, we determine the coefficient of repeatability (CR) for the given parameters.

In **chapter five** we will evaluate the effect of pursed lips breathing (PLB) in patients with severe-to-very-severe COPD (GOLD stages 3 and 4) on inspiratory parameters: (forced inspiratory volume in first second (FIV1), IC, maximal inspiratory flow at 50% of VC (MIF50), and peak inspiratory flow (PIF)) and secondary outcome parameters including FEV1, forced VC (FVC), oxygen saturation, end-tidal CO2 tension (ET-CO2), BF, and dyspnea

In **chapter six** we present our results of short-acting bronchodilators and their effect on FEV1 and inspiratory parameters, where we also consider the number of patients responding for these parameters beyond the random variation in patients with COPD.

In **chapter seven** we will explore whether protocol-driven test ordering reduces the number of redundant pulmonary function tests, decreases the number of outpatient visits, and increases the cost effectiveness of patient work-up in comparison to physician-driven test ordering.

In **chapter eight** we will summarize the results, discuss further research in this field and make final conclusions so far.

Reference List

1. From the World Health Oganisation; the definition of COPD. Available from http://www.who.int/respiratory/copd/definition/en/. 2011. 7-5-2011.

2. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease:ATS statement. *Am J Respir Crit Care Med* 1995, 152.

3. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2009. Available from: http://www.goldcopd.org.

4. From the Global Strategy for Asthma Mangement and Pprevention, Global initiative for Asthma (GINA). Available from http://www.ginasthma.com. 2010.

5. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R *et al*.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26: 948-968.

6. Standards for the diagnosis and treatment of patients with chronic obstructive pulmonary disease. 2004.Available from: http://www.thoracic.org/clinical/copd-guidelines/resources/copddoc.pdf

7. 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: 1117-1122.

8. Kerstjens HA: The GOLD classification has not advanced understanding of COPD. *Am J Respir Crit Care Med* 2004, 170: 212-213.

9. Celli BR, Cote CG, Marin JM, Casanova C, Montes de OM, Mendez RA *et al*.: The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 2004, 350: 1005-1012.

10. Celli BR, MacNee W: Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004, 23: 932-946.

11. Karoli NA, Rebrov AP: [The BODE index as a predictor of unfavourable prognosis in chronic obstructive pulmonary disease (by the results of a prospective study)]. *Ter Arkh* 2007, 79: 11-14.

12. Curtis JR: Palliative and end-of-life care for patients with severe COPD. *Eur Respir J* 2008, 32: 796-803.

13. Archibald CJ, Guidotti TL: Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab* 1987, 1: 45.

14. Cazzola M: Should the choice of a long-acting bronchodilator in the long-term therapy of chronic obstructive pulmonary disease depend entirely on the onset of action? *Respiration* 2006, 73: 410-411.

15. Guyatt GH, Townsend M, Nogradi S, Pugsley SO, Keller JL, Newhouse MT: Acute response to bronchodilator. An imperfect guide for bronchodilator therapy in chronic airflow limitation. *Arch Intern Med* 1988, 148: 1949-1952.

16. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162: 216-220.

17. O'Donnell DE, Lam M, Webb KA: Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999, 160: 542-549.

18. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1: 307-310.

19. Nici L, Donner C, Wouters E, Zuwallack R, Ambrosino N, Bourbeau J *et al*.: American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation. *Am J Respir Crit Care Med* 2006, 173: 1390-1413.

20. Troosters T, Casaburi R, Gosselink R, Decramer M: Pulmonary rehabilitation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005, 172: 19-38.

21. Breslin EH: The pattern of respiratory muscle recruitment during pursed-lip breathing. *Chest* 1992, 101: 75-78.

22. Fagevik OM, Westerdahl E: Positive expiratory pressure in patients with chronic obstructive pulmonary disease--a systematic review. *Respiration* 2009, 77: 110-118.

23. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Grazzini M *et al*.: Chest wall kinematics and breathlessness during pursed-lip breathing in patients with COPD. *Chest* 2004, 125: 459-465.

24. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Binazzi B *et al*.: Patterns of chest wall kinematics during volitional pursed-lip breathing in COPD at rest. *Respir Med* 2007, 101: 1412-1418.

25. Garrod R, Dallimore K, Cook J, Davies V, Quade K: An evaluation of the acute impact of pursed lips breathing on walking distance in nonspontaneous pursed lips breathing chronic obstructive pulmonary disease patients. *Chron Respir Dis* 2005, 2: 67-72.

2

Recommendations for the measurement of FIV1 values in COPD

Frank J. Visser, Sunil Ramlal, P.N. Richard Dekhuijzen, Yvonne F. Heijdra

Respiration. 2008;76(1):46-52. Epub 2007 Nov 21.
PMID: 18032882

Abstract

Background: In contrast to static inspiratory parameters such as vital capacity and inspiratory capacity, information on forced inspiratory volume in 1 s (FIV_1) in patients with chronic obstructive pulmonary disease (COPD) is limited. Objectives: It was the aim of this study to investigate the influence of the preceding expiratory manoeuvre and the optimal number of manoeuvres on FIV_1 values.

Methods: In 169 patients with COPD, FIV_1 manoeuvres were performed after a forced (FIV_1-Fe) and a slow (FIV_1-Se) expiration. To investigate the optimal number of the FIV_1-Se manoeuvres, 8 attempts were performed.

Results: The variability of FIV_1-Fe was greater than that of FIV_1-Se. The mean difference between FIV_1-Se and FIV_1-Fe was 0.21 litres ($p < 0.01$) and dependent on the Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage. The higher the GOLD stage, the greater the difference between the 2 techniques. The correlation coefficient between FIV_1-Se and FIV_1-Fe was high ($r = 0.89$, $p = 0.01$), but there was a poor agreement between these parameters (limits of agreement -0.52 to 0.94 litres). Five manoeuvres were needed to obtain an optimal FIV_1-Se. There was no association with the GOLD stage.

Conclusions: In COPD patients, FIV_1-Se are less variable than FIV_1-Fe, the agreement between the 2 manoeuvres is poor, and at least 5 FIV_1-Se manoeuvres are needed to get an acceptable FIV_1. This holds for all GOLD stages.

Introduction

The severity of chronic obstructive pulmonary disease (COPD) is defined by the degree of expiratory airflow limitation. It is essential for the diagnosis and provides a useful description of the severity of pathological changes in COPD [1]. However, the impact of COPD on an individual patient depends not just on the degree of airway limitation but also on the severity of symptoms and complications of the disease. It is well known that the correlation between the subjective improvements in dyspnoea and increases in forced expiratory volume in 1 s (FEV_1) after inhalation of bronchodilators is poor [2-5].

However, in patients with COPD and expiratory flow limitation at rest, changes in inspiratory and forced vital capacities (FVC) after bronchodilator use may be an objective tool for prescribing bronchodilators, even in the absence of a significant increase in FEV_1 [5].

This was shown by Taube et al. [4] who found that in patients with severe COPD (mean FEV_1 38% of predicted), the reduction in dyspnoea after inhalation of a β_2-adrenoreceptor agonist was closely correlated with the change in parameters of forced inspiration, and particularly forced inspiratory volume in 1 s (FIV_1), but not with changes in parameters of forced expiration or hyperinflation including inspiratory capacity (IC).

In contrast with this study, Richter et al. [6] did not find significant improvement in FIV_1 and IC after tiotropium or formoterol, but they commented that this study was not powered to find significant differences.

Newton et al. [7] characterized bronchodilator effect in terms of flow and volume response. Flow response was determined according to changes in FEV_1. The volume response was ascertained by examining the bronchodilator effect on IC, residual volume and FVC. Overall, flow response occurred in 33% of patients in the severe hyperinflation group (total lung capacity >133% of predicted value) and in 26% of patients in the moderate hyperinflation group (115% < total lung capacity < 133% of predicted value), and volume response occurred in up to 76% of patients in the severe hyperinflation group and in 62% of patients in the moderate hyperinflation group.

In the recently published American Thoracic Society/European Respiratory Society (ATS/ERS) statement on clinical pulmonary function testing [8], no recommendations are made for the measurement of inspriratory parameters

including FIV_1. Therefore, it is unclear how FIV_1 should be measured. One of the important issues is to define to what extend the expiratory manoeuvre preceding forced inspiration influences FIV_1. Also, the number of manoeuvres needed to adequately measure FIV_1 is not specified [9-11]. Taube et al. [4] used a slow expiration before the measurement of FIV_1, whereas most commercially available lung function apparatuses provide FIV_1 after a fast expiration.

Therefore, in the present study, we aimed to formulate recommendations for the measurement of the FIV_1 manoeuvre in COPD patients by focusing on the following research items.

The difference in FIV_1 when performed after a slow (FIV_1-Se) or a forced (FEV_1-Fe) expiration and its relation to the severity of COPD as defined by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage.

The correlation between and interchangeability of FIV_1-Se and FIV_1-Fe.

The variability of FIV_1-Se and FIV_1-Fe and its dependence on the severity of COPD.

The optimal number of FIV_1-Se manoeuvres in order to get an acceptable value for FIV_1, and whether this number is dependent on the severity of COPD.

Material and Methods

A total of 169 consecutive patients (121 males) who met the GOLD criteria [1, 12] for COPD were recruited from our outpatient clinic. Inclusion criteria were reversibility of FEV_1 <12% of predicted normal value and <200 ml [11], age ≥40 years, smoker or former smoker (≥10 pack-years), stable disease, and able to perform lung function tests. Patients on oral corticosteroids or antibiotics in the month before inclusion, patients with symptomatic heart failure, respiratory diseases other than COPD, a history of asthma, allergic rhinitis and active cancer disease (except basal cell carcinoma of the skin) were excluded. The study was approved by the Hospital Medical Ethical Committee and all patients gave informed consent.

Study Design

Each patient was requested to discontinue all bronchodilating medications starting at least 24 h before execution of the tests. Before the tests, the 3.00-litre calibration syringe was used at 3 different emptying and filling speeds to check linearity as recommended by the ATS and ERS standards.

The ambient (room) temperature was measured before each test session to adequately perform body temperature, pressure and saturation corrections on the flows and volumes. In random order, the following FIV_1-Fe and FIV_1-Se techniques were performed.

To measure FIV_1-Fe, patients performed as many manoeuvres as needed (with a maximum of 8) to achieve 3 adequate and acceptable flow volume curves, according to conventional ATS/ERS criteria [8]. By this technique, the FIV_1 is obtained after fast expiration (FIV_1-Fe). For FIV_1-Se, 8 maximal forced inspirations after a slow and maximal expiration were obtained (FIV_1-Se). Maximal inspiration was obtained when a plateau was reached, or after at least 8 s duration of the inspiration.

If during the inspiratory manoeuvres VC was reached before FIV_1, then FIV_1 = VC. The largest FVC, FEV_1 and FIV_1 are recorded. For the predicted FEV_1 and FVC, the normal values of the European Community for Steel and Coal were used [11].

To make a fair comparison between the results of FIV_1-Fe versus FIV_1-Se, the first 3 adequate FIV_1-Fe manoeuvres were compared with the first 3 adequate FIV_1-Se manoeuvres. The flow-volume curves were measured with the V-MAX20 (Sensor Medics, ViaSys, Conshohocken, Pa., USA).

In order to obtain proper inspiratory parameters after a slow expiration, we started the measurement during slow expiration and stopped the procedure when the patient reached forced inspiratory vital capacity (FIVC), otherwise the software of the V-MAX20 rejects the obtained values.

Data Analysis

Data are presented as mean ± SD, or in the case of a non-normal distribution, as median and interquartile range. Differences between FIV_1-Fe and FIV_1-Se were analyzed with the paired t test. One-way ANOVA and post-hoc multiple comparison t test (Bonferroni) were used to test differences between the GOLD classes. The correlation between FIV_1-Fe and FIV_1-Se was calculated by Pearson's correlation test. The interchangeability of FIV_1-Fe and FIV_1-Se was investigated as described by Bland and Altman [13]. We defined the variability of FIV_1-Se, FIV_1-Fe and FEV_1 as the difference of the best minus the second best of the first 3 adequate attempts and expressed this difference as the percentage of its average. The differences of the means between these groups were tested with a paired sample t test. To assess the optimal number of FIV_1 manoeuvres after a slow expiration, the best of 3, 4, 5 and 6 attempts was compared with the best of 8 attempts. This

was compared with and plotted against the percentage of failures during these attempts. After obtaining the optimal number of FIV_1 manoeuvres, the effect of the GOLD stage on this optimal number in comparison with 8 manoeuvres was examined (Spearman rank correlation). All tests used a 2-tailed significance level of 5%, unless otherwise stated. SPSS version 11 (2001) was used to analyse the data.

Results

In table 1, the demographic characteristics of our patients are presented.
Of the 169 patients, 1 failed to perform 3 adequate forced expiratory flow curves, 16 patients failed to perform 3 adequate inspiratory flow curves after forced expiration, and 6 patients failed to perform 3 adequate inspiratory flow curves after slow expiration. The difference of 6 versus 16 failed manoeuvres of FIV_1-Se versus FIV_1-Fe was significant ($p < 0.05$, χ^2). In 168 patients, FEV_1 was 1.52 ± 0.71 litres, in 153 patients, FIV_1-Fe was 2.5 ± 0.79 litres, and in 163 patients, FIV_1-Se was 2.71 ± 0.77 litres. In 79 of 153 subjects who performed 3 FIV_1-Fe manoeuvres, the FIVC was substituted for the FIV_1. This was also done in 71 of 163 subjects who performed 3 FIV_1-Se manoeuvres.

Table 1 *Demographic characteristics*

Subjects, n	169
Age, years	65±9.9
Males	121 (72)
Females	48 (28)
Height, cm	171±8
Weight, kg	76.5±16.9
FEV_1, % predicted	64.74±18.96
GOLD 1	30 (17.4)
GOLD 2	61 (35.5)
GOLD 3	44 (25.6)
GOLD 4	34 (19.8)

Data are presented as number of patients, with percentage in parenthesis, or mean ± SD.

Difference in FIV_1 when Performed after FIV_1-Se or FEV_1-Fe and Its Relation to the Severity of COPD as Defined by the GOLD Stage

Of 169 patients, 151 completed both 3 adequate FIV_1-Se and 3 adequate FIV_1-Fe manoeuvres. FIV_1-Se was greater than FIV_1-Fe by 0.21 litres ($p < 0.001$, t test). The results of the 3 FIV_1-Fe attempts and the best of the first 3 FIV_1-Se attempts and their mean difference are shown in table 2

Table 2. *The best of 3 attempts of FIV1-Se and FIV1-Fe and their difference*

	Patients, n	Mean, litres	SD
Best FIV_1-Fe	151	2.50	0.79
Best FIV_1-Se	151	2.71	0.82
Difference between FIV_1-Se and FIV_1-Fe	151	0.21	0.36

The difference between FIV_1-Se and FIV_1-Fe is dependent on the GOLD stage.

Figure 1 presents the difference between FIV_1-Se and FIV_1-Fe according to the GOLD stage and shows that the difference for COPD stage 1 was 4%, which rose to 16% for GOLD stage 4. The ANOVA between group test showed a significant ($p = 0.017$) difference for the relative differences of the means. The post-hoc multiple comparison t test (Bonferroni) showed that the mean difference in GOLD stage 4 differs significantly from GOLD classes 1 and 2.

Correlation between and Interchangeability of FIV_1-Se and FIV_1-Fe

The Pearson correlation coefficient between FIV_1-Se and FIV_1-Fe was 0.89 ($p < 0.01$). However, FIV_1 values obtained with the FIV_1-Se and FIV_1-Fe manoeuvres were not interchangeable. The mean difference between FIV_1-Se and FIV_1-Fe was 0.21 litres, and the limits of agreement were -0.52 to 0.94 litres. In figure 2, we present the difference between the 2 methods plotted against the average of the difference as described by Bland and Altman [13].

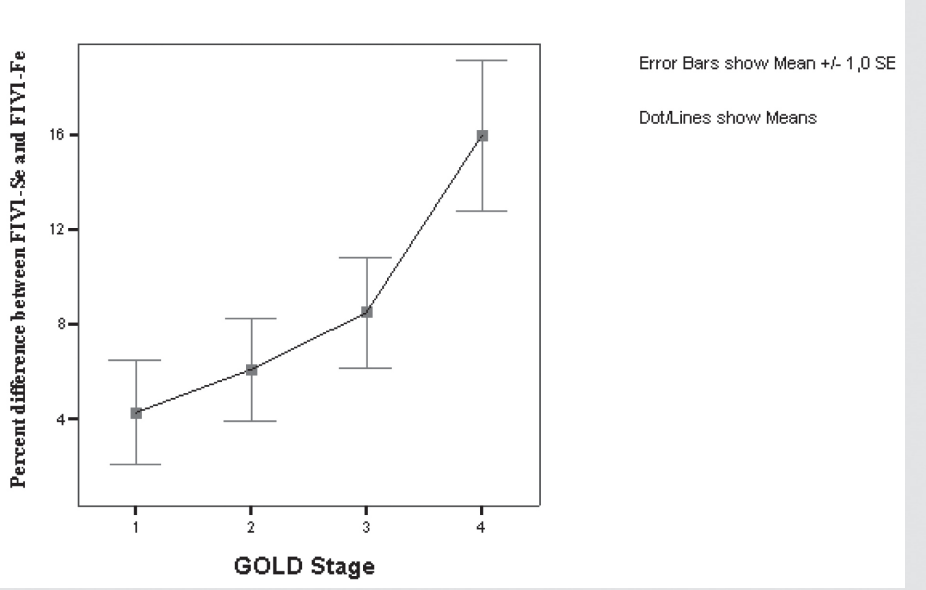


Figure 1. *Difference of FIV_1-Se and FIV_1-Fe with respect to the severity of COPD. The difference is expressed as percentage of the average value on the vertical axis versus the GOLD stage on the horizontal axis. Error bars show means ± 1.0 SE; the line indicates the mean.*

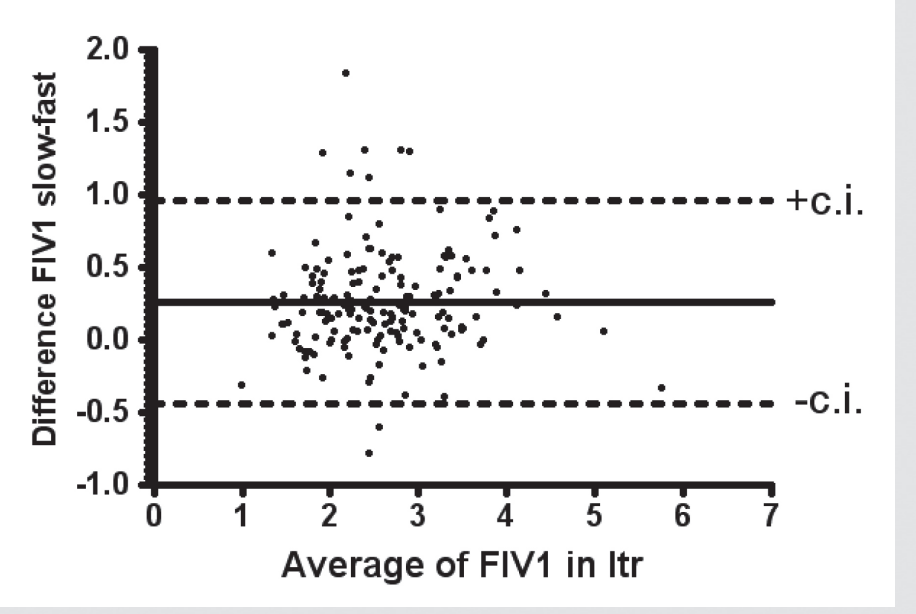


Figure 2. *Interchangeability of FIV_1-Se and FIV_1-Fe. On the x-axis, the mean FIV_1 in liters is given, on the y-axis, the difference between FIV_1-Se and FIV_1-Fe. Horizontal lines represent the means and the 95% confidence intervals (CI).*

Variability of FIV_1-Se and FIV_1-Fe and Its Dependence on the Severity of COPD

FEV_1 measurements were far less variable than FIV_1-Fe and FIV_1-Se (paired sample t test, $p < 0.001$), as can be seen in table 3.

Table 3. *Variability of FIV1-Fe, FIV1-Se and FEV1 measurements*

	Patients n	Median %	Mean %	IQR %	90th percentile %
Variability of FIV_1 after fast expiration	153	4.90	7.56	7.44	14.71
Variability of FIV_1 after slow expiration	163	3.83	5.25	5.96	11.62
Variability of FEV_1	162	2.06	2.92	2.86	07.05

The variability between the 3 adequate manoeuvres expressed as the best minus the second best as percentage of their mean. IQR = Interquartile range.

FIV_1-Fe showed a higher variability than FIV_1-Se (paired sample t test of the difference, $p = 0.011$). To test if the variability is related to the GOLD class we did an ANOVA between group test for FIV_1-Se ($p = 0.532$) and FIV_1-Fe ($p = 0.448$). We found that variability was not related to COPD stage (table 4).

Table 4. *Variability according to GOLD class*

GOLD class	Patients n	Mean %	SEM %	Median %	IQR %
Variability FIV_1-Se					
All	163	5.25	0.36	3.83	5.96
1	29	6.23	0.86	4.73	6.65
2	58	5.01	0.58	3.75	5.05
3	43	4.70	0.71	3.24	3.94
4	33	5.53	0.86	4.35	7.27
Variability FIV_1-Fe					
All	153	7.56	0.83	4.90	7.44
1	28	6.71	1.90	3.36	5.6
2	56	6.79	1.31	5.19	7.08
3	40	7.21	1.17	4.44	7.97
4	29	10.34	2.63	7.58	8.79

Values are expressed as percentage of the mean of the best and second best manoeuvre. SEM = Standard error of the mean; IQR = interquartile range.

Optimal Number of FIV_1-Se Manoeuvres in Order to Get an Acceptable Value for FIV_1, and whether This Number Is Dependent on the Severity of COPD

Because FIV_1-Se was higher and less variable than FIV_1-Fe, we searched for the optimal number of FIV_1-Se manoeuvres in order to get an acceptable value for FIV_1-Se and investigated whether this optimal number depends on the GOLD stage.

The greater the number of attempts, the closer the FIV_1-Se value was to that obtained from the best of 8 attempts (the difference is expressed as a percentage of the best of 8) (fig. 3, line a, right vertical axis). However, the greater the number of attempts, the more patients failed to perform 8 adequate tests (fig. 3, line b, left vertical axis). We wanted to find the number of attempts that gives a good compromise between the most accurate results and the maximum number of patients to successfully complete the test (fig. 3).

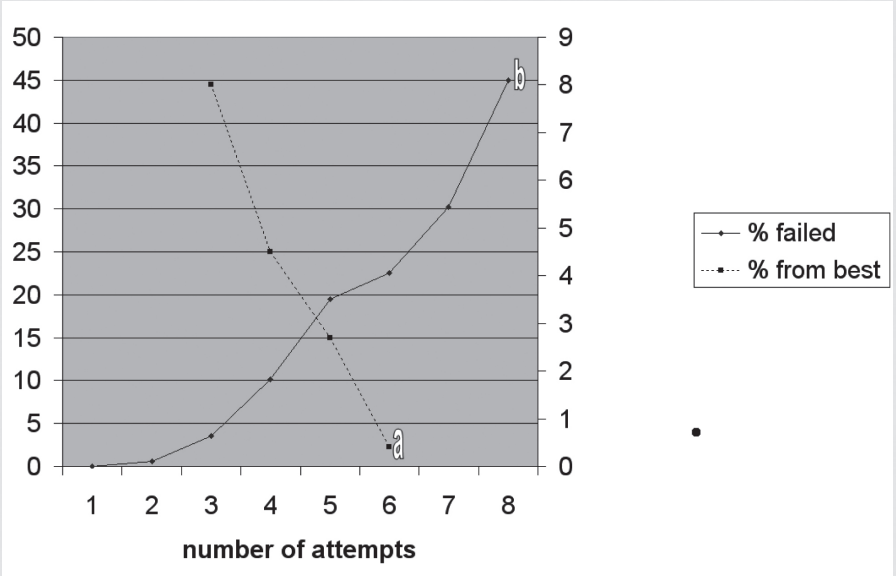


Figure 3. *Percentage of failed FIV_1 manoeuvres and the difference from the best attempt with respect to the number of attempts. Graphical representation of the percentage of failure of adequate inspiratory manoeuvres (line b, left vertical axis) and the percentage of difference of FIV_1 from the best of 8 attempts versus the number of adequate attempts (line a, right vertical axis).*

We found that with 5 adequate attempts, we had a deviation from the best of 8 attempts of 6.4% at the 95th percentile and 2.7% at the 90th percentile, so 10% of our COPD population had a difference of 2.7% or more, and 5% of our population a difference of 6.4% or more from the best of 8 attempts. Of our COPD patients, 80.5% completed 5 or more adequate FIV_1 manoeuvres out of 8 attempts.
We analysed the results of 5 attempts (the optimal number) with respect to the COPD stage (I--IV) of the patients and found no association between the number of inadequate efforts and the stage of disease (Spearman correlation, $r = 0.007$, $p = 0.931$).

Discussion

This study shows that in patients with COPD, the highest FIV_1 values were measured after a slow expiration. The correlation between FIV_1-Se and FIV_1-Fe was high, but the limits of agreement were wide. Therefore, these manoeuvres are not interchangeable. The variability of FIV_1-Fe was higher than that of FIV_1-Se, and more patients failed to perform 3 adequate FIV_1-Fe manoeuvres than 3 adequate FIV_1-Se manoeuvres. Five adequate FIV_1-Se attempts were necessary to get an optimal FIV_1 value in all GOLD classes.
In this study, we have shown that the mean difference between FIV_1-Se and FIV_1-Fe was 0.21 litres. No other published comparison of the 2 methods for measuring FIV_1 has been found. Taube et al. [4] have used FIV_1 after a slow expiration for their study; however, they did not mention the reason for choosing this method. The reason for the difference between the 2 methods might be the airway collapse associated with a forced expiration [14]. This increases the physiological dead space leading to an increased end-expiratory volume (or residual volume). The beginning of the following inspiration is at a higher volume, therefore resulting in less room and a less advantageous length-tension relationship of the inspiratory muscles, resulting in a smaller FIV_1 after a forced expiration. In our opinion, it is possible that after a forced expiration, the airway resistance increases which also causes a lower FIV_1.
Ewald et al. [15] found that forced inspiratory capacity was significantly larger after a slow expiration than after an antecedent forced expiration by 170 ml and he thus recommended a slow expiration but did not mention FIV_1 or the variability of the IC.

In this study, we provide evidence that the mean difference between FIV_1-Se and FIV_1-Fe was dependent on the GOLD stage; the lower the FEV_1, and thus the higher the GOLD stage, the greater the difference between FIV_1-Se and FIV_1-Fe. This is what we expected: the more severe the COPD, the greater the airway collapses after forced expiration, and therefore, the greater the difference between these manoeuvres. We have not found this published in literature before.

The present study shows that the values of FIV_1-Se and FIV_1-Fe were not interchangeable because of the poor limits of agreement between these manoeuvres (mean difference 0.21 litres, limits of agreement -0.52 to 0.94 litres). We found that of 9 outliers (as seen in the Bland-Altman plot), 7 were in the higher range. This might be because of very collapsible airways. The 2 outliers on the lower side might have upper airway collapsibility as seen for instance in patients with obstructive sleep apnoea syndrome.

Despite the poor agreement, we found a good correlation between FIV_1-Se and FIV_1-Fe (Pearson, $r = 0.89$, $p < 0.01$). This is not surprising, because we are dealing with the same parameter, namely FIV_1.

Our results indicate that the variability of FIV_1-Se was significantly smaller (paired sample t test, $p = 0.011$) and had less dispersion than that of the FIV_1-Fe (median 3.8 vs. 4.9%). This finding is important, and therefore, we are recommending to use FIV_1-Se as the method of choice. In addition, we found that patients failed more often in performing adequate FIV_1-Fe than FIV_1-Se manoeuvres. FEV_1 shows less variability (median 2.06%) than FIV_1-Se and FIV_1-Fe. We think this is because of the check valve phenomenon during forced expiration, which explains the effort-independent part of the expiration [14, 16--18]. In contrast, inspiration is fully effort dependent, and thus, shows variations in patient effort [14]. Other possible causes of variability may be that the manoeuvre is less known to both the patients and the pulmonary function assistants, as is the way FIV_1-Se has to be performed on the V-MAX20 series.

As far as we know, no articles are available in literature dealing with the reproducibility of FIV_1. However, we did find a study of Taube et al. [19] who measured reproducibility as coefficients of variation, not within 1 measurement, but after bronchodilators on 3 different days. They found a coefficient of variation of 43 and 45% for FEV_1 and FIV_1, respectively, after fenoterol in 13 patients with COPD. However, it is interesting that the

coefficients of variation derived from FEV_1 and FIV_1 are close to each other. Probably, this rather shows the day-to-day variability of airway obstruction in COPD patients.

In our study, the variability did not depend on the GOLD class, and the ANOVA between group tests for FIV_1-Se and FIV_1-Fe were not significant. This is important because FIV_1-Se remains a valid reproducible test in all COPD stages. We did not find this conclusion in previous literature.

The assessment of an optimal number of FIV_1 manoeuvres is complicated: the greater the number of manoeuvres, the more the patient tires. In patients with COPD, more than 8 manoeuvres might be exhausting, especially when the patient also has to make forced expiratory manoeuvres in order to get important parameters like FEV_1. It seems to be difficult to get more than 5 adequate manoeuvres out of 8 attempts (more than 20% of COPD patients failed; fig. 3, line b).

The difference of the best of 8 versus the best of 3, 4, 5 or 6 inspiratory manoeuvres is higher, the lower the number of attempts compared (fig. 3, line a). Therefore, we searched for the number of attempts that compromises between tiring of the patient and the highest value (fig. 3). Consequently, we recommend 5 manoeuvres of FIV_1 after a slow expiration, which will not tire the patients and provide an acceptable difference between 5 and 8 manoeuvres (fig. 3).

The variability of FIV_1 is higher than that of FEV_1. The FEV_1 manoeuvre has to be repeated just 3 times if the best minus the second best is <5% of the best or <100 ml.

There is no association (Spearman correlation 0.007, $p = 0.931$) between the number of failed manoeuvres and the Gold stage, so the amount of manoeuvres needed is not depended on the GOLD stage.

To better measure FIVC and FIV_1-Se, the software of many spirometers will have to be updated. Almost all of them simply store and print the results of FIVC which immediately followed the 'best' FVC manoeuvre (or in some cases, the best 3 FVC manoeuvres).

In a following study, using the above-mentioned method, we will focus on the intraday and interday repeatability of inspiratory parameters and the reaction to several bronchodilators.

Conclusions

In COPD patients, FIV_1 values after a slow exhalation are higher and more stable than after a forced exhalation, FIV_1-Se and FIV_1-Fe are not interchangeable, more patients failed to perform 3 adequate FIV_1-Fe than FIV_1-Se manoeuvres, and at least 5 adequate FIV_1-Se manoeuvres are needed in order to get an acceptable FIV_1 (2.7% difference at the 90th percentile and 6.5% at the 95th percentile). This recommendation holds for all 4 GOLD stages.

Acknowledgements

We would like to thank Bert Brakert and Bard Kelder, lung function technicians, for their excellent support. We appreciate the assistance of Gareth Collier and Roney de Abreu for methodical and linguistic assistance with the manuscript.

Reference List

1. COPD: 2003 update. *Eur Respir J* 2003;22:1--2.

2. Archibald CJ, Guidotti TL: Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab* 1987;1:45.

3. Guyatt GH, Townsend M, Nogradi S, Pugsley SO, Keller JL, Newhouse MT: Acute response to bronchodilator. An imperfect guide for bronchodilator therapy in chronic airflow limitation. *Arch Intern Med* 1988;148:1949--1952.

4. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000;162:216--220.

5. Cazzola M: Should the choice of a long-acting bronchodilator in the long-term therapy of chronic obstructive pulmonary disease depend entirely on the onset of action? *Respiration* 2006;73:410--411.

6. Richter K, Stenglein S, Mucke M, Sieder C, Schmidtmann S, Harnest U, *et al*: Onset and duration of action of formoterol and tiotropium in patients with moderate to severe COPD. *Respiration* 2006;73:414--419.

7. Newton MF, O'Donnell DE, Forkert L: Response of lung volumes to inhaled salbutamol in a large population of patients with severe hyperinflation. Chest 2002;121:1042--1050.

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

9. Lung function testing: selection of reference values and interpretative strategies: official statement ATS. *Am Rev Respir Dis* 1991;144:1202--1216.

10. Standardization of spirometry, 1994 update. American Thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107--1136.

11. Quanjer PH, Tammeling GJ: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993;6(suppl 16):5--40.

12. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease: ATS statement. *Am J Respir Crit Care Med* 1995;152:S78--S121.

13. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307--310.

14. Gibson GJ: Clinical Tests of Respiratory Function, ed 2. *London, Chapman & Hall,* 1996.

15. Ewald FW Jr, Tenholder MF, Waller RF: Analysis of the inspiratory flow-volume curve. Should it always precede the forced expiratory maneuver? *Chest* 1994;106:814--818.

16. Fry DL, Hyatt RE: Pulmonary mechanics: a unified analysis of the relationship between pressure, volume and gas flow in the lungs of normal and diseased subjects. *Am J Med* 1960;29:672--689.

17. Mead J, Turner JM, Macklem PT, Litle JB: Significance of the relationship between lung recoil and maximum expiratory flow. *J Appl Physiol* 1967;22:95--108.

18. Pride NB, Permutt S, Riley RL, Romberger-Barnea B: Determinants of maximal expiratory flow from the lungs. *J Appl Physiol* 1967;23:646--662.

19. Taube C, Kanniess F, Gronke L, Richter K, Mucke M, Paasch K, et al: Reproducibility of forced inspiratory and expiratory volumes after bronchodilation in patients with COPD or asthma. *Respir Med* 2003;97:568--577.

3

The one hour test-retest repeatability of FEV1 in patients with COPD

Frank J Visser, Sunil Ramlal, Ben Pelzer,
P N Richard Dekhuijzen, Yvonne F Heijdra

Submitted

Abstract

Background: In patients with chronic obstructive pulmonary disease (COPD), a significant bronchodilator response must exceed the random variation (RV). Therefore, it is important that the type of scatter is homoscedastic, thereby reducing the chance of under or overestimating the RV for low or high parameter values. The aim of this study was to investigate the RV (type and quantity) of the forced expiratory volume in one second (FEV1) measurement.

Methods: spirometry was performed in 79 stable COPD patients. The FEV1 was measured five times in one day. FEV1 values taken within one hour were compared. The coefficient of repeatability (CR) was calculated and linear regression was performed to investigate the type of scatter (i.e., homo- or heteroscedastic).

Results: the type of scatter was heteroscedastic for the FEV1 when the difference was expressed as an absolute value. However, when the difference was expressed as the percent change from the initial value or predicted value, we found homoscedastic scatter. The CR within one hour of the FEV1 was 12% when expressed as the percent change from the initial value.

Conclusions: to obtain a more homoscedastic scatter, the percent change in FEV1 is a more appropriate measure than the absolute change.

Introduction

The severity of chronic obstructive pulmonary disease (COPD) is defined by the degree of expiratory airflow limitation [1]. After using bronchodilators, many COPD patients may experience less dyspnea but do not show significant reversibility in their FEV1, which is defined by a 12% improvement from the initial value and at least 200 milliliters in volume [2]. However, there is no clear consensus as to how to express the reversibility of the FEV1 [2]. The three most commonly described methods are the percent change from the initial FEV1 value, the absolute change in this parameter's value and the percent change from normal predicted values [2].

For an individual patient with COPD, a significant bronchodilator response must exceed the random variation for the parameter of interest. Therefore, it is important to know the type of variation or scatter that exists for the parameter of interest. Figure 1 shows theoretical datasets for a test-retest FEV1 with different types of scatter. One dataset shows a 200 milliliter difference between the test and retest and the other dataset shows a difference of 12% from the initial value between the test and retest.

It is important to know the type of scatter for the parameter of interest. Typically, homoscedastic scatter is desired. The consequence of the ATS-ERS 200 milliliter criterion as stated by Pellegrino et al., [2] dictates that for a patient with COPD and a FEV1 of 0.5 liter, an improvement of at least 40% from the initial value is necessary before one may conclude that this is beyond random variation (Figure 1, panel B).

The importance of this 200 milliliter improvement was also raised by Hansen [3] and illustrated by the publication of Han et al., [4], which examined the prevalence of bronchoreversibility in subjects enrolled in the National Emphysema Trial. This later study found that only 22% of patients met the ATS-ERS criterion at least once (of 7 time points) and 10.2% of patients met the criterion at the first time point. However, if the standard of only exhibiting a 12% increase from the basal value was used, 452 (83%) of subjects met this criterion at least once [4].

The literature establishing this 200 milliliter criterion has to be scrutinized (see discussion section) because none of the referenced literature used the recommended method described by Bland and Altman [5].

The primary aim of this study was to investigate the type of scatter that best applies to the observed (absolute and percent) changes in FEV1. In addition, we sought to determine the coefficient of repeatability (CR) for the FEV1 [2,5].

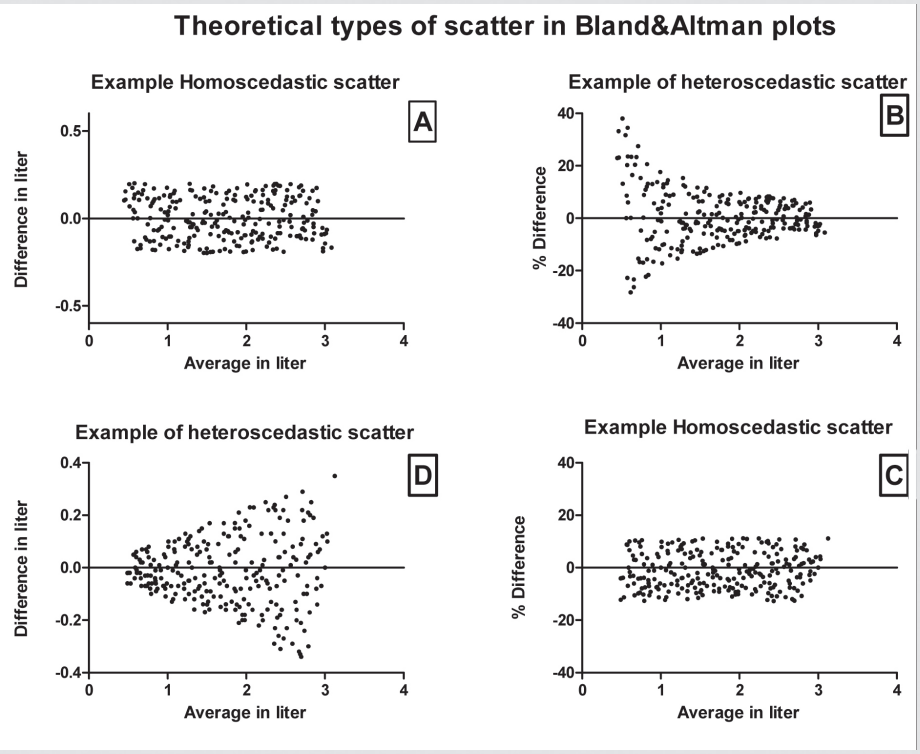


Figure 1. *Bland and Altman plots with theoretical types of scatter. The vertical axis shows the test (T) retest (R) difference (T-R); the horizontal axis shows the average (T+R)/2. Each point corresponds with one test-retest pair.*
Panel A: Data set with a random scatter that is not dependent on the parameter value (homoscedastic). We used 200 milliliter between test and retest as recommended for the FEV1.
Panel B: Same dataset, but now the percent change from the average parameter value is shown. The scatter is highly dependent on the parameter value. In panel A, the scatter can be described with just one value, e.g., the coefficient of repeatability (CR). In panel B, the scatter cannot be precisely described with one value because the scatter on the low and high parameter values will not precisely reflect a fixed CR (heteroscedastic). Thus, for describing random variation, a homoscedastic scatter is preferred.
Panel C: Another dataset but now with a homoscedastic scatter of 12% from the average value.
Panel D: Represents the scatter of the same dataset but expressed as the absolute difference.

Materials and methods

A total of 79 (58 male) consecutive patients who met ATS-ERS [6] criteria for COPD were recruited from our outpatient clinic. These were the same patients we selected for a study on inspiratory parameters by Visser et al., [7]. The criteria for inclusion were the following: patient age ≥ 40 years, a smoker or former smoker (≥10 years), presentation of a stable disease state and an ability to perform lung function tests. Excluded patients were those on oral corticosteroids or antibiotics in the month before the study's initiation or those who had symptomatic heart failure, respiratory diseases other than COPD, a history of asthma, allergic rhinitis or active cancer (except for basal cell carcinoma of the skin). The study was approved by the Hospital Medical Ethical Committee and all patients gave their informed consent.

Methods

Patients were asked not to use short-term bronchodilators for six to eight hours prior to the study and long-term bronchodilators for at least 12 hours before the study. Tiotropium and theophylline (b.i.d.) were not allowed for 24 hours prior to the spirometric test.

Lung function tests were performed five times on the first day (at 9, 10, 11 A.M. and 2 and 3 hours P.M.) and once at 9 hours A.M. over the following two weeks. Between these two days, the medication did not change. On the second day, these patients were requested to discontinue bronchodilator use as was done on day one.

For expiratory parameters, three suitable flow volume curves were produced in accordance with conventional ATS/ERS criteria [8]. The largest forced vital capacity (FVC) and FEV1 were recorded. Of the retests, only the FEV1 was recorded. For the predicted FEV1 and FVC, the normal values of the European Respiratory Society were used [9].

The flow–volume curves were measured with a V-MAX20 (Sensor Medics, ViaSys, Conshohocken, PA, USA).

Analysis

The five intra-day lung function parameter data were analyzed with a repeated measures ANOVA followed Bonferroni's multiple comparison tests.

The type of scatter (homoscedastic or heteroscedastic) was determined by the method described by Visser et al., [7].
The coefficient of repeatability (CR), as observed for the one hour, intra-day and inter-day time points, was determined by the method described by Bland and Altman [5]. A one-tailed test was utilized to calculate the lower limits of normal subjects [9].

Results

Seventy-nine patients were included for day one (intra-day measurements) and 76 were measured again within the following two weeks. For two patients, we were unable to have an appointment within the two week period. In addition, in one patient, there was an exacerbation. The baseline patient characteristics are summarized in Table 1.

***Table 1**. Characteristics of the COPD patient group*

No. of subjects	79	**Range**
Male, %	73	
Height, m, mean (SD)	1.702 (0.1)	1.49-1.94
Weight, kg, mean (SD)	75.4 (15.9)	39.7-128
BMI, kg/m^2, mean (SD)	25.7 (4.96)	18-47
Age, yrs, mean (SD)	65.4 (8.69)	44-83
Current smoker, %	59	
FEV1, mean (SD)	1.48 (0.70)	0.37-2.65
Predicted FEV1, mean (SD)	2.83 (0.60)	1,26-4.23
FEV1, % predicted, (SD)	48.7 (12.6)	13.5-79.8
GOLD 1 number	13	
GOLD 2 number	24	
GOLD 3 number	28	
GOLD 4 number	14	

COPD= Chronic Obstructive Pulmonary Disease; BMI= Body Mass Index; FEV1=Forced Expiratory Flow in one second; GOLD= Global Initiative for Chronic Obstructive Lung Disease (stage 1, 2, 3 or 4).

The intra-day mean and SD values did not differ significantly (repeated measures ANOVA) on different occasions for this day (see Table 2). Therefore, we took three pairs of measurements for each patient with a one hour difference (9–10, 10–11 and 14–15 hours) for determination of the type of scatter and the one hour coefficient of repeatability.

Table 2. *Mean and (SD) values of the FEV1 on five occasions during the day*

Parameter/time	9	10	11	14	15
FEV1	1.48 (0.70)	1.48 (0.69)	1.48 (0.90)	1.48 (0.70)	1.41 (0.68)

FEV1=Forced Expiratory Flow in one second.

The type of scatter for the FEV1

The scatter of the one hour differences in FEV1 versus the average FEV1 value across the two time periods is shown in Figure 2,

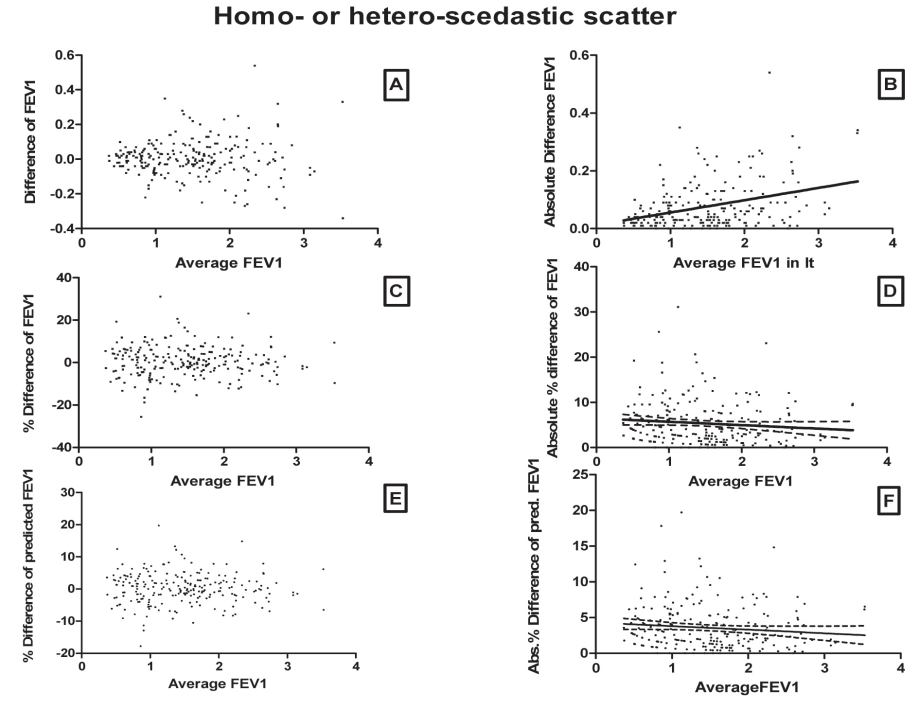


Figure 2. *Panel A: Bland–Altman plot showing the absolute difference of FEV1 versus the average, which shows a heteroscedastic scatter of the FEV1.*
Panel B: Absolute difference of FEV1 versus average FEV1 plus linear regression, which shows a slope in the regression line.
Panel C: Difference FEV1 now presented as the percentage of the average value, which shows a homoscedastic scatter.
Panel D: Percentage difference of FEV1 versus average FEV1 plus linear regression, which shows a flat slope of the regression line, which is not significantly different from zero.
Panel E: Difference of FEV1 as percentage of predicted FEV1.
Panel F: Absolute difference of FEV1 as percentage of predicted FEV1 plus linear regression line showing a homoscedastic scatter.
The dashed lines in panel B and D and F represent the confidence interval of the regression line.

Panel A. The scatter becomes wider when the average FEV1 increases. Panel B presents the same data except that negative difference values are made positive and a linear regression line is added to the figure. The regression line is not flat (P<0.0001), which would be the case if the scatter was independent of the FEV1 (Table 3).
On the other hand, when the difference in FEV1 is expressed as a percentage of the average FEV1 (Figure 2, panel C) or as a percentage of the predicted FEV1 (Figure 2, panel E), an evenly distributed scatter can be seen along the whole range of the average FEV1. The slope of the linear regression line is now nearly flat (Figure 2, panels D and F) and is not significantly different from zero (Table 3).

Table 3. *Linear regression: standardized slopes (r) and P values; tests whether the slope is significantly different from zero N=237*

Parameter	Standardized slope (r)	P value (significance from zero slope)
FEV1	0.37	<0.0001
FEV1% initial	0.10	0.1 (NS)
FEV1%pred	0.11	0.1 (NS)

FEV1=Forced Expiratory Flow in one second; FEV1% initial=FEV1 as percentage form initial value; FEV1%pred=FEV1 as percentage form predicted value. See also Figures 2 and 3 and the text for further explanation.

The random variability presented by the coefficient of repeatability

The coefficients of repeatability for FEV1 are graphically presented in the Bland and Altman plots in Figure 3. The spread around the solid line (no difference) can be seen and the coefficient of repeatability is presented by the dotted line with ±1.64 × standard deviation. The spread appears to be relatively even, thereby indicating homoscedascity.

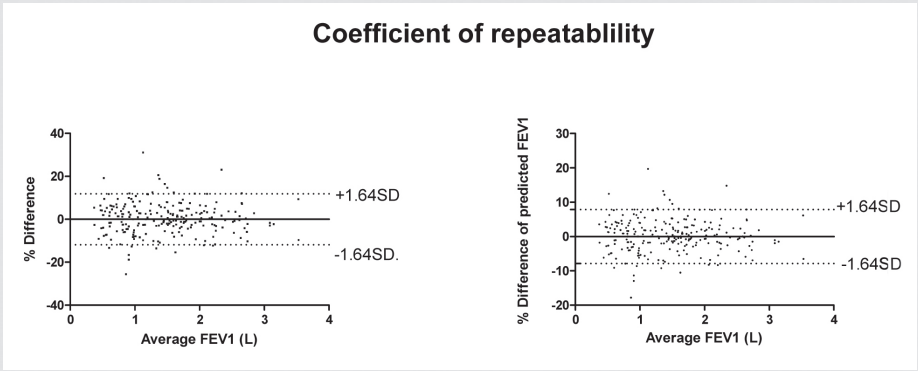


Figure 3. *Left panel: Bland–Altman plot of the percentage difference versus the average of the FEV1. The coefficient of repeatability (CR) corresponds with the dotted lines (e.g., the coefficient of repeatability is 11.8%, which is 1.64x SD). Right panel: Blant-Altman plot of the percentage difference of the FEV1 to the normal predicted value versus the average FEV1. The coefficient of repeatability is 7.8%.*

In Table 4, the results of the coefficients of repeatability (CR) are summarized. The FEV1 presented as a percentage of the predicted value shows a coefficient of repeatability of 7.8% (Figure 3). In addition, the one hour coefficients of repeatability, the 5 hour and the between day coefficients of repeatability are summarized for the FEV1 percent change from initial values (Table 4).

Table 4. *Coefficients of repeatability (CR) retested after one hour (n=237), after five hours (n=79) and after 3-8 days (n-76)*

Parameter	Retest after one hour	Retest after five hours	Retest another day
FEV1 % initial	11.8%	13.7%	17.9%

FEV1=Forced Expiratory Flow in one second; FEV1% initial=FEV1 as percentage form initial value; FEV1%pred=FEV1 as percentage form predicted value.

Discussion

Type of scatter

This study shows that within the same subject, the one hour difference in FEV1 can be more appropriately expressed as the percentage of the initial value or the percentage of the predicted value rather than merely as the difference between these two values. Notably, both percentage methods produce a more homoscedastic scatter. The measured FEV1 showed a scatter that was significantly dependent on the average parameter value (Figure 2, panels A and B and Table 3). Consequently, these data appear heteroscedastic if we present these differences as liters. On the other hand, if we represent the difference as the percentage of the average value (which was similar to the initial value) or as the percentage of the predicted value, we found no significant dependence on the FEV1 value.

Several studies have addressed the variability of lung function parameters, which is especially evident with forced expiration [10-14]. However, these studies used the variation coefficient instead of the method described by Bland and Altman [5]. Therefore, the type of spread was not investigated. According to the ATS-ERS, the method of Bland and Altman is described as the preferred method for investigating random variation and this method allows the type of spread to be readily visible [2,5].

Studies mentioned by Pellegrino et al., [2] are described in the following text and abstracted in Table 5.

References made to studies with measurements of bronchodilator responses in normal and healthy subjects are not relevant to the context of random variation among patients with COPD because they neither described the population of interest nor the random variation [15-17]. However, the study of Dales at al., [16] is worth mentioning because they found that among 2609 healthy subjects with an FEV1 greater than 80% of the predicted value the upper limit of improvement (to terbutaline) was 9% of the predicted normal value. Therefore, more than a 9% improvement of the predicted normal value after bronchodilator administration can be seen as an indication of reversibility in this group. However, examining random variation in patients with COPD in this publication was not done in this publication.

Pellegrino et al., [18] focused on the partial forced expiratory flow volume curves after bronchodilators in 78 patients, 50 asthmatics and 28 COPD patients. In that study, only 26 subjects were tested with placebo and the

number of asthmatics and COPD patients that were included was not reported. Of these placebo tested patients, the authors mentioned the mean test retest difference (1%) and the SD (7%). However, they did not use the method described by Bland and Altman and did not report the type of scatter.

Anthonissen et al., [10] tested 985 patients with COPD after bronchodilators. There were no test-retest or placebo groups used in this study so there were no data examining random variation. Therefore, the remarks on the type of scatter were not relevant for the assessment of random variation.

Sourk et al., [14] tested 40 subjects with placebo and 72 with bronchodilators. Of the 40 placebo tested patients, they presented the confidence intervals of the test–retest differences as a percentage of the initial value and as an absolute difference. However, they did not investigate the type of scatter.

Tweeddale et al., [19] tested 150 patients with obstructive lung diseases (asthma and COPD), which included a test-retest after 20 minutes. They did not use the method described by Bland and Altman but divided the obstructed patients into the three following subgroups by their average FEV1: group A, 0.5-1.1 liter; group B, 1.15-2.4 liter; group C, greater than 2.45 liter. They found significantly different values for the SD of the test-retest differences when expressed as a percentage of the predicted value (Levine test) and no significant difference was found between these groups when the absolute difference was taken into account. Thus, their results are in contrast with our findings. However, they used a different method without visualizing the test-retest scatter, a different patient population and a different time interval. In addition, 150 data points were examined in this previous, while 237 data points were included in our study.

Enright et al., [20] described the short-time repeatability of FEV1 but they examined this variable only with one prebronchodilator test and took the difference between the highest and second highest FEV1 value. Furthermore, only 9.5% (170 subjects) of their study population was mentioned to have COPD (and 11% had asthma) and the rest of the patient population was not described. They also did not perform a test on the type of scatter (homo or heteroscedastic), which is of the utmost importance as we have described in this article.

Table 5. *Abstract of references used for the latest official ATS-ERS recommendation for improvement of FEV1 after bronchodilators*

Reference	Diagnosis	Number of subjects	Test-retest or placebo	Used method for type of scatter
Dales et al. [16]	Healthy subjects	2609	None, dilator was tested	None
Pellegrino et al. [18]	Asthma/COPD	26	Placebo	None
Anthonissen et al. [10]	COPD	958	None, only dilators were tested	
Sourk et al. [14]	Obstructive?	40	placebo	None
Tweeddale et al. [17]	Asthma/COPD	150	Test-retest	Levine test

Thus, none of these previous studies used the recommended Bland and Altman method for describing test-retest results for random variation and the only study that investigated the type of scatter (by the Levine test) used a mixed patient population. Therefore, we believe that our study provides the strongest evidence that the type of random scatter for FEV1 is best revealed if the test values are expressed as a percent change among patients with COPD. Consequently, we think that the recommendation of at least a 200 milliliter absolute difference [2] may be reconsidered.

Although we are reporting statistically significant improvements, the clinical relevance is difficult to establish because of the weak correlation between dyspnea and FEV1 changes after bronchodilators. If we repeatedly find no significant improvement within a subject, we may be observing a placebo effect.

The practical consequence of this finding is that more patients with serious COPD can be considered responders to bronchodilators when we use the criterion of 12% of the initial value (or 8% of the predicted value) and omit the 200 milliliter absolute response standard.

Random variability

The one hour repeatability was by far the most important for random variability because most interventions of interest (such as bronchodilator response) can be measured within one hour. Subjects must exceed this level

of random variation before it can be decided that an improvement of an intervention can be attributed to that intervention.

We decided to pool our patient data (with one hour differences measured at three time points a day per patient) to obtain more data points and, thus, more reliable results. This pooling was possible because we found no significant differences between the group means and spreads as well as no significant differences in the parameters for measurements between 9 a.m. and 2 p.m. This result may be in contrast with Calverley et al., [21] and van Noord et al., [22] who found significantly lower values at 3 and 6 a.m.; however, we did not perform measurements at these hours.

We chose the CR instead of the more popular variation coefficient because it more precisely reflects repeatability and provides a graphical representation of the type of scatter, which has been stated by Bland and Altman [5] and adheres to the recommendations of the ERS-ATS committee [2].

The **intra-day** coefficients of repeatability are important to know when we are performing interventions that take more than one hour, i.e., medications such as theophylline, tiotropium or other interventions that take more time to retest.

We selected the 9 to 14 hour difference because the FEV1 group's mean did not change during this interval. There was a small but significant decrease in FEV1 during the 9 to 15 hour interval (Table 2); therefore, we took the 9 to 14 interval as our difference. We think that this decrease may be due to the fact that patients at the end become tired of repeating the procedure five times a day, which is compounded in the meantime by not being allowed to take any bronchodilator drugs; moreover, there may have been some circadian effect [21,22]. The higher CR value, compared to the one hour value, is expected due to the larger time interval.

The **inter-day random variability** improvements of interventions taking more than one day can be considered as beyond random variation when the above-mentioned improvements are taken into account. Between the two days, patients did not change their medication and no exacerbations occurred. Another supporting point of this inter-day CR was that we were unable to see all of the patients on an exact inter-day interval of one week because we were limited by when our patients were able to visit our outpatient department. As the smallest interval was three days and the largest interval was eight days, analyses were all conducted within two weeks. In general, the longer the interval between the two measurements (from one hour to

several days), the higher the CR obtained. This result may be caused by the longer time period due to weather or other environmental factors affecting patient accessibility.

Limitations of this study

Subjects in this study included only patients with COPD, so our findings do not extend to normal patients or those with asthma or restrictive diseases. In addition, the random variation within these groups may have been different. We used all data (including outliers) to construct the Bland and Altman plots; however, small sample sizes can influence the linear regression analysis of the transformed Bland and Altman plots when the outliers are in the lower or upper zones of the average parameter value.

Conclusions

Random variation of FEV1 is less dependent on changing parameter values when presented as a percentage of the initial value or the predicted value than as the absolute difference.

In individual patients with COPD, their improvement after intervention must exceed the given coefficients of repeatability.

We think the 200 milliliter criterion as stated by the ERS may be reconsidered.

Reference List

1. Fabbri LM, Hurd SS: Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update. *Eur Respir J* 2003, 22: 1-2.

2. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R *et al*.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26: 948-968.

3. Hansen JE: Assessment of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 36: 972.

4. Han MK, Wise R, Mumford J, Sciurba F, Criner GJ, Curtis JL *et al*.: Prevalence and clinical correlates of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 35: 1048-1056.
5. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1: 307-310.

6. Celli BR, MacNee W: Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004, 23: 932-946.

7. Visser FJ, Ramlal S, Pelzer B, Dekhuijzen PN, Heijdra YF: Random variation of inspiratory lung function parameters in patients with COPD: a diagnostic accuracy study. *BMC Pulm Med* 2010, 10: 28.

8. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A *et al*.: Standardisation of spirometry. *Eur Respir J* 2005, 26: 319-338.

9. Quanjer Ph.H., Tammeling G.J.: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993, 6: 5-40.

10. Anthonisen NR, Wright EC: Bronchodilator response in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1986, 133: 814-819.

11. Dompeling E, van Schayck CP, Molema J, Akkermans R, Folgering H, van Grunsven PM *et al*.: A comparison of six different ways of expressing the bronchodilating response in asthma and COPD; reproducibility and dependence of prebronchodilator FEV1. *Eur Respir J* 1992, 5: 975-981.

12. Enright PL, Connett JE, Kanner RE, Johnson LR, Lee WW: Spirometry in the Lung Health Study: II. Determinants of short-term intraindividual variability. *Am J Respir Crit Care Med* 1995, 151: 406-411.

13. Nickerson BG, Lemen RJ, Gerdes CB, Wegmann MJ, Robertson G: Within-subject variability and per cent change for significance of spirometry in normal subjects and in patients with cystic fibrosis. *Am Rev Respir Dis* 1980, 122: 859-866.

14. Sourk RL, Nugent KM: Bronchodilator testing: confidence intervals derived from placebo inhalations. *Am Rev Respir Dis* 1983, 128: 153-157.

15. Lorber DB, Kaltenborn W, Burrows B: Responses to isoproterenol in a general population sample. *Am Rev Respir Dis* 1978, 118: 855-861.

16. Dales RE, Spitzer WO, Tousignant P, Schechter M, Suissa S: Clinical interpretation of airway response to a bronchodilator. Epidemiologic considerations. *Am Rev Respir Dis* 1988, 138: 317-320.

17. Watanabe S, Renzetti AD, Jr., Begin R, Bigler AH: Airway responsiveness to a bronchodilator aerosol. I. Normal human subjects. *Am Rev Respir Dis* 1974, 109: 530-537.

18. Pellegrino R, Rodarte JR, Brusasco V: Assessing the reversibility of airway obstruction. *Chest* 1998, 114: 1607-1612.

19. Tweeddale PM, Alexander F, McHardy GJ: Short term variability in FEV1 and bronchodilator responsiveness in patients with obstructive ventilatory defects. *Thorax* 1987, 42: 487-490.

20. Enright PL, Beck KC, Sherrill DL: Repeatability of spirometry in 18,000 adult patients. *Am J Respir Crit Care Med* 2004, 169: 235-238.

21. Calverley PM, Lee A, Towse L, van Noord J, Witek TJ, Kelsen S: Effect of tiotropium bromide on circadian variation in airflow limitation in chronic obstructive pulmonary disease. *Thorax* 2003, 58: 855-860.

22. van Noord JA, Aumann JL, Janssens E, Verhaert J, Smeets JJ, Mueller A *et al*.: Effects of tiotropium with and without formoterol on airflow obstruction and resting hyperinflation in patients with COPD. *Chest* 2006, 129: 509-517.

4

Random variation of inspiratory lung function parameters in patients with COPD: a diagnostic accuracy study

Frank J Visser, Sunil Ramlal, Ben Pelzer
P N Richard Dekhuijzen, Yvonne F Heijdra

BMC Pulm Med. 2010 May 14;10:28.
PMID: 20470374

Abstract

Background: In chronic obstructive pulmonary disease (COPD), the response of the forced expiratory volume in 1 second (FEV1) after bronchodilator application is weak. Inspiratory parameters like the forced inspiratory volume in 1 second (FIV1) and inspiratory capacity (IC) can be responsive to bronchodilators. In an individual patient with COPD, a significant bronchodilator response must at least exceed the random variation for that parameter. Therefore, it is important that the type of scatter is homoscedastic, as the chance of underestimating or overestimating the random variation for low or high parameter values is minimized. The aim of this study is to investigate the random variation (type and quantity) of inspiratory parameters.

Methods: In 79 stable COPD patients, spirometry was performed.
The forced inspiratory volume in 1 second (FIV1), inspiratory capacity (IC), maximal inspiratory flow at 50% (MIF50) and peak inspiratory flow (PIF) were measured five times in one day and again within two weeks of the first measurement. The values of these parameters, taken within one hour, within one day and between two different days, were compared. The coefficient of repeatability (CR) was calculated, and, in addition, linear regression was performed to investigate the type of scatter (homo- or heteroscedastic) of the measured parameters.

Results: The type of scatter was heteroscedastic for all of the parameters when the differences were expressed as absolute values; however, when the differences were expressed as the percent change from the initial values, we found a more homoscedastic scatter. The CR within one hour of each parameter expressed as the percent change from the initial value was: IC, 19%; FIV1, 14%; PIF, 18%; MEF50, 21%.

Conclusions: To obtain a more homoscedastic scatter, percentage changes in FIV1, IC and MIF50 are more appropriate than absolute changes.
In an individual patient with COPD, a significant improvement for a particular parameter must at least exceed the above-mentioned CR.

Background

The severity of chronic obstructive pulmonary disease (COPD) is defined by the degree of expiratory airflow limitation. It is essential for diagnosis and provides a useful description of the severity of pathological changes in COPD [1]. It is, however, well known that the correlation between the subjective improvements in dyspnea and the increases in Forced Expiratory Volume in 1 second (FEV1) after inhalation of bronchodilators is low [2-4]. Many COPD patients do not show significant reversibility of FEV1 after bronchodilators, as defined by a 12% improvement from the initial value and at least 200 ml [5], but may experience less dyspnea from their use. Taube and co-workers [4] demonstrated that this change in dyspnea may be related to improvements in inspiratory flow rates. These authors found that in patients with severe COPD (FEV1 mean was 38% of the predicted normal value), the reduction in dyspnea after the inhalation of a beta (2)-adrenoreceptor agonist was closely correlated to the change in parameters of forced inspiration, particularly for the forced inspiratory volume in 1 second (FIV1), but not with changes in parameters of forced expiration. They also concluded that, "In less severe COPD or asthma, the reduction in dyspnea was associated with the improvements in both FIV1 and FEV1, but in severe COPD with the improvement in FIV1 only" [6].
O'Donnell et al. found a correlation between the change of the Inspiratory Capacity (IC) after bronchodilator administration, dyspnea and duration of exercise [7,8]. In 2005, a published ATS/ERS statement on clinical pulmonary function testing [9] made no recommendations on the measurement of inspiratory parameters including FIV1 . Therefore, it is unclear how FIV1 and other inspiratory parameters should be measured and which improvements in a patient are beyond random variation for these parameters after the use of bronchodilators or other interventions.
How FIV1 should be measured in patients with COPD was the subject of a previous study by our group. We found that the optimal FIV1 was obtained immediately after a slow expiration (in contrast to a forced expiration) and that at least five forced inspiratory maneuvers should be performed [10].
However, there is no clear consensus about how to express reversibility in subjects with airflow limitation [5]. The two most commonly described methods are the percent change of the initial value and the absolute change in the parameter value. As the percent change from the initial value is too sensitive at very low values, as measured in severe obstructive patients, a

third method uses the percent change from the predicted normal value [5]. For the inspiratory parameters under study, no accepted predicted normal values are available; hence, we used the first two methods.

For an individual patient with COPD, a significant bronchodilator response must at least exceed the random variation for the parameter of interest. Therefore, it is important to know which type of variation or scatter exists for that parameter. Figure 1 shows a theoretical dataset of a test-retest lung-function parameter with different types of scatter. In the left panel, we made the amount of scatter the same for each value of the parameter, called "homoscedastic" scatter. For the whole range of the parameter, we can use the same value for the random variation, and a difference of more than 0.2 is beyond the random variation.

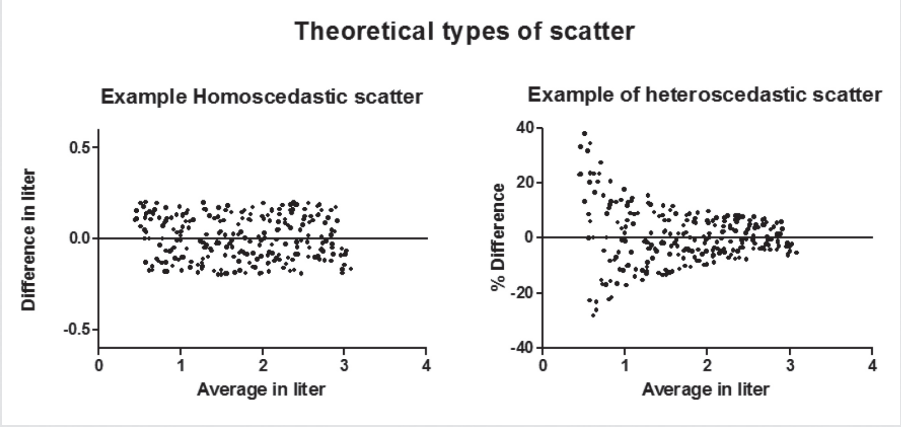


Figure 1. *Bland and Altman plots with theoretical types of scatter. The vertical axis shows the test (T) retest (R) difference (T-R); the horizontal axis shows the average (T+R/2). Each point corresponds with one test-retest pair. Left panel: dataset with a random scatter not dependent on the parameter value.*
Right panel: same dataset, but now the percent change from the average parameter value is shown; the scatter is highly dependent on the parameter value. In the left panel, the scatter can be described with just one value, e.g., the coefficient of repeatability (CR). In the right panel, the scatter cannot be precisely described with one value because the scatter on the low and high parameter values will not precisely reflect a fixed CR. Thus, for describing random variation, a homoscedastic scatter is preferred.

The same dataset is used in the right panel, in which the differences are related to the (average) parameter value (percent difference) but now the amount of scatter depends on the parameter value: the higher the value, the less (in this example) scatter or random variation there is; this type of scatter is called "heteroscedastic". Therefore, it is important to know the type of scatter of the parameters in which we are interested. A more homoscedastic scatter is desired when we express the differences as absolute differences or as relative to the parameter value.
The first topic of this study is to investigate which type of scatter applies to the (absolute and percent) changes in inspiratory parameters (FIV1, Inspiratory Capacity (IC), Peak Inspiratory Flow (PIF) and Maximal Inspiratory Flow at 50% (MIF50)). Next, we determine the coefficient of repeatability (CR) for the given parameters [5,11].

Methods

A total of 79 (58 male) consecutive patients who met ATS-ERS [12] criteria for COPD were recruited from our outpatient clinic. Criteria for inclusion were a patient age ≥ 40 years, a smoker or former smoker (≥10 pack years), stable disease and an ability to perform lung function tests. Excluded patients were those on oral corticosteroids or antibiotics in the month before inclusion, or those who had symptomatic heart failure, respiratory diseases other than COPD, a history of asthma, allergic rhinitis or active cancer disease (except basal cell carcinoma of the skin). The study was approved by the Hospital Medical Ethical Committee, and all patients gave informed consent.
Patients were asked not to use short-term bronchodilators for the six to eight hours prior to the study and long-term bronchodilators for at least 12 hours before the study. Tiotropium and theophylline b.i.d. were not allowed to be used for the 24 hours prior to the spirometric test.
Before the tests, a 3.00-liter calibration syringe was used at three different emptying and filling speeds to check linearity, as recommended by ATS and ERS standards [9]. The ambient (room) temperature was measured before each test session so that BTPS corrections on the flows and volumes were adequately performed.
Lung function tests were performed five times on the first day (9, 10, 11, 14 and 15 hours) and once at nine hours within the following two weeks.

Between the two days, the medication did not change. Also, on the second day the patients were requested to discontinue bronchodilators as on day one. For expiratory parameters, three adequate and acceptable flow volume curves were produced in accordance with conventional ATS/ERS criteria [9]. The largest forced vital capacity (FVC) and FEV1 were recorded. For the predicted FEV1 and FVC, the normal values of the European Respiratory Society were used [13].
For inspiratory parameters, five adequate IC measurements and maximal forced inspirations after a slow and maximal expiration were obtained. Full inspiration was obtained when a plateau in the flow was reached or after at least an eight-second duration of the inspiration. Of these five maneuvers, we took the highest value obtained for the FIV1, IC, PIF and MIF50 [10]. IC was measured by the method described by Hadcroft and Calverly [14] immediately before each forced inhalation.
If, during the inspiratory maneuvers, the vital capacity (VC) was reached before the FIV1, then FIV1=VC. The flow–volume curves were measured with a V-MAX20 (Sensor Medics, ViaSys, Conshohocken, PA, USA).
In order to obtain proper inspiratory parameters after a slow expiration, we began the measurement during the slow expiration and stopped the procedure when the patient reached maximal inspiration; otherwise, the V-MAX20 software rejects the values obtained.

Analysis
The five intra-day lung function parameter data were analyzed with the repeated measures ANOVA and Bonferroni's multiple comparison tests.
The type of scatter (homoscedastic or heteroscedastic) was determined as follows. The differences between each test and retest value pair versus the average value were plotted as described by Bland and Altman [11]. Negative differences were transformed to positive values by taking the absolute values of the differences. We applied linear regression of these transformed differences on the average value of the parameter. When there is a pure homoscedastic scatter, the regression line will be close to horizontal and the slope will not significantly differ from zero. When there is a heteroscedastic scatter, the slope of the regression line will be significantly different from zero. For each parameter, scatter plots were made for both absolute differences and percentage differences from the average value. With linear regression, we tested for the significances of the slopes. Instead of slopes,

we present standardized slopes, i.e., correlation coefficients, as these can be compared across parameters and methods.

The coefficient of repeatability (CR), as established within one hour, intra-day and inter-day, was determined by the method described by Bland and Altman [11]. The CR was determined as 1.64 times the standard deviation of the differences, represented as absolute values or as percentages of the average values. We performed a one-tailed test instead of a two-tailed test, as the interventions we were interested in were expected to improve a parameter; hence, we took 1.64 times SD instead of 1.96 X SD. This use of a one-tailed test is analogous to the way in which lower limits of normals are calculated [13]. CR was used instead of the more common coefficient of variation (CV) because CV does not take into account the type of scatter that can be visualized by the scatter plots of Bland and Altman [11]

Results

Seventy-nine patients were included for day one (intra-day measurements), and 76 were measured again within the following two weeks. For two patients, we were unable to get an appointment within the two weeks, and in one there was an exacerbation. The baseline patient characteristics are summarized in Table 1.
The intra-day mean and SD values did not differ significantly (repeated measures ANOVA) on different occasions that day (see Table 2). Therefore, we took three value pairs per parameter for each patient with a one-hour difference (9–10, 10–11 and 14–15 hours) for determination of the type of scatter and the one-hour coefficient of repeatability.

The type of scatter for the inspiratory parameters FIV1, IC, PIF and MIF:
The scatter of differences in IC values versus average IC value on two occasions in-between one hour is shown in Figure 2, panel A. The scatter becomes wider when the average IC increases. Panel B presents the same data, except that negative difference values are made positive and a linear regression line is added to the figure. The regression line is not flat ($P<0.0001$), as it would be if the scatter was independent of the IC (Table 3)

Table 1. *Characteristics of the COPD patient group.*

Number of subjects	79	Range
Male (%)	58 (73)	
Height, m, mean (SD)	1.702 (0.1)	1.49-1.94
Weight, kg, mean (SD)	75.4 (15.9)	39.7-128
BMI, kg/m^2, mean (SD)	25.7 (4.96)	18-47
Age, yrs, mean (SD)	65.4 (8.69)	44-83
Current smoker (%)	47 (59)	
FEV1, mean (SD)	1.48 (0.70)	0.37-2.65
Predicted FEV1, mean (SD)	2.83 (0.60)	1,26-4.23
FEV1, % predicted, (SD)	48.7 (12.6)	13.5-79.8
GOLD 1 number (%)	13 (17)	
GOLD 2 number (%)	24 (30)	
GOLD 3 number (%)	28 (35)	
GOLD 4 number (%)	14 (18)	

COPD= Chronic Obstructive Pulmonary Disease; BMI=Body Mass Index; FEV1=Forced Expiratory Flow in One second; GOLD= Global Initiative for Chronic Obstructive Lung Disease (stage 1,2,3 or 4) .

Table 2. *Mean and (SD) values of lung function parameters on five occasions during the day*

Parameter/time	9	10	11	14	15
FEV1	1.48 (0.70)	1.48 (0.69)	1.48 (0.90)	1.48 (0.70)	1.41 (0.68)
FIV1	2.70 (0.85)	2.71 (0.87)	2.69 (0.87)	2.65 (0.84)	2.62 (0.81)
IC	2.13 (0.69)	2.15 (0.68)	2.17 (0.72)	2.18 (0.75)	2.14 (0.79)
MIF50	4.68 (1.63)	4.61 (1.61)	4.65 (1.62)	4.51 (1.55)	4.47 (1.18)
PIF	4.96 (1.67)	4.92 (1.66)	4.95 (1.67)	4.89 (1.61)	4.84 (1.65)

FEV1=Forced Expiratory Flow in one second; FIV1= Forced Inspiratory Flow in One second; IC- Inspiratory Capacity; MIF50= Maximal Inspiratory Flow at 50% ; PIF= peak Inspiratory Flow.

On the other hand, when the difference in IC is expressed as a percentage of the average IC (Figure 2, panel C), an evenly distributed scatter can be seen along the whole range of the average IC. The slope of the linear regression line is now nearly flat (Figure 2, panel D) and is not significantly different from zero (Table 3).

Table 3. *Linear regression: standardized slopes (r) and P values; tests whether the slope is significantly different from zero. N=237*

Parameter	Standardized slope (r)	P value (significance from zero slope)
FEV1	0.37	<0.0001
%FEV1	0.10	0.1 (NS)
IC	0.27	< 0.0001
%IC	0.05	0.46 (NS)
FIV1	0.20	0.002
%FIV1	0.13	0.045
PIF	0.20	0.002
%PIF	0.24	0.0002
MIF50	0.22	0.0009
%MIF50	0.14	0.0267

FEV1=Forced Expiratory Flow in one second; FIV1= Forced Inspiratory Flow in One second; IC- Inspiratory Capacity; MIF50= Maximal Inspiratory Flow at 50% ; PIF= peak Inspiratory Flow. See also Figures 2 and 3 and the text for further explanation.

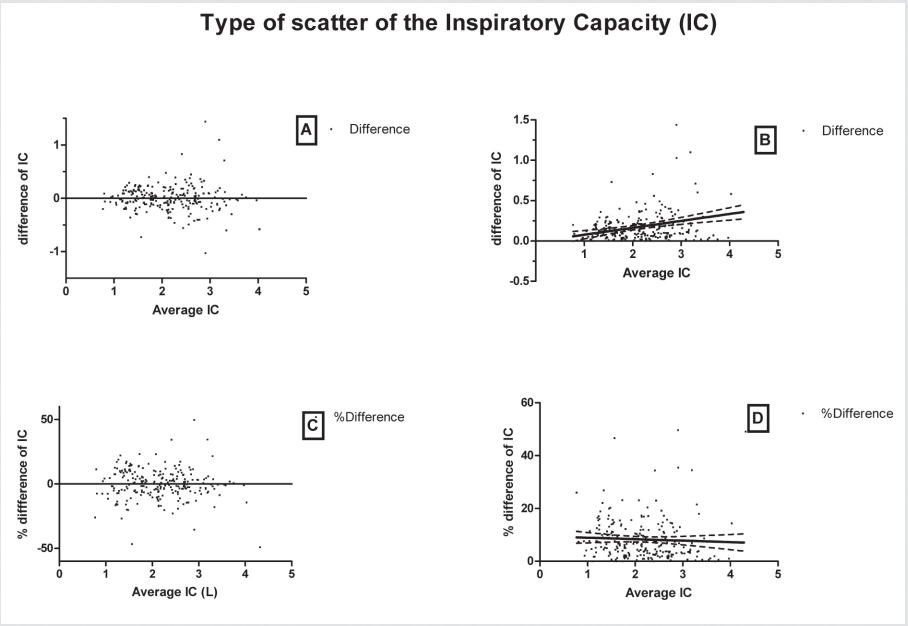


Figure 2. *Panel A: Bland–Altman plot showing the absolute difference of Inspiratory Capacity (IC) versus the average; it shows a heteroscedastic scatter of the IC. Panel B: Absolute difference of IC versus average IC plus linear regression; this panel shows a slope in the regression line. Panel C: Difference IC now presented as the percentage of the average value; this panel shows a homoscedastic scatter. Panel D: Percentage difference of IC versus average IC plus linear regression; this panel shows a flat slope of the regression line, not significantly different from zero. The dashed lines in panel B and D represent the confidence interval of the regression line.*

The other parameters were investigated in the same way as the IC, and the results of the linear regression are presented in Table 3. To visualize the (more homoscedastic) spread, the scatter plots of the FIV1, IC, MIF50 and PIF expressed as percentage differences can be seen in Figure 3.

We did not find a flat regression line for either presentation of the parameters (as differences in liters or as percentage differences of the average value); however, for all but the PIF, we found a more flat regression line corresponding to a lower (r) value when the percentage difference of the average value was used (less significant difference from zero, as can be seen in Table 3).

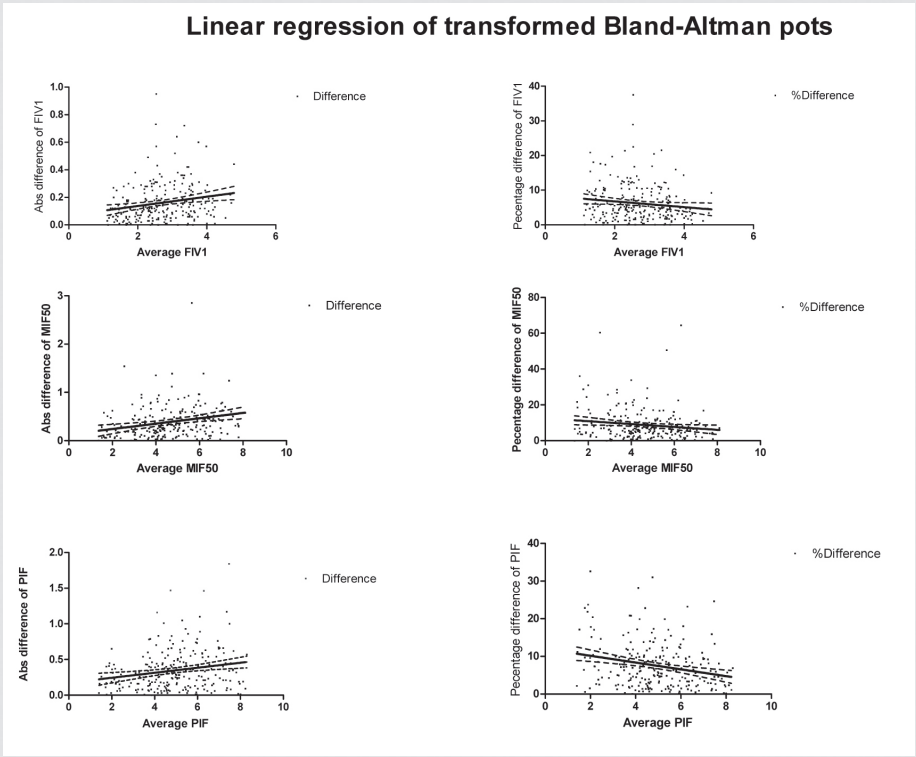


Figure 3. *Left panels show the scatter (absolute differences and regression lines for FIV1, MIF50 and PIF). Right panels show the percentage differences for the same parameters. The dashed lines represent the confidence intervals of the regression lines. All slopes of the regression lines are significantly deviated from flat (zero); however, apart from the PIF, the percentage differences show regression lines that are more flat (closer to a zero slope). FIV1= Forced Inspiratory Flow in One second; MIF50= Maximal Inspiratory Flow at 50% ; PIF= peak Inspiratory Flow*

The random variability presented by the coefficient of repeatability

The coefficients of repeatability for IC, FIV1, MIF50 and PIF are graphically presented as Bland–Altman plots in Figure 4, panels A–D. The spread around no difference (solid line) can be seen, and the coefficient of repeatability is presented as the dotted lines ±1.64 × standard deviation. We can also see the relatively even spread around the solid line, which is an indication of a more homoscedastic spread. In patients with COPD, we found that the one-hour random variabilities expressed as the coefficients of repeatability (CR) for the lung function parameters are: IC: 19%, FIV1: 14%, PIF: 18% and MEF50: 21% (Table 4).

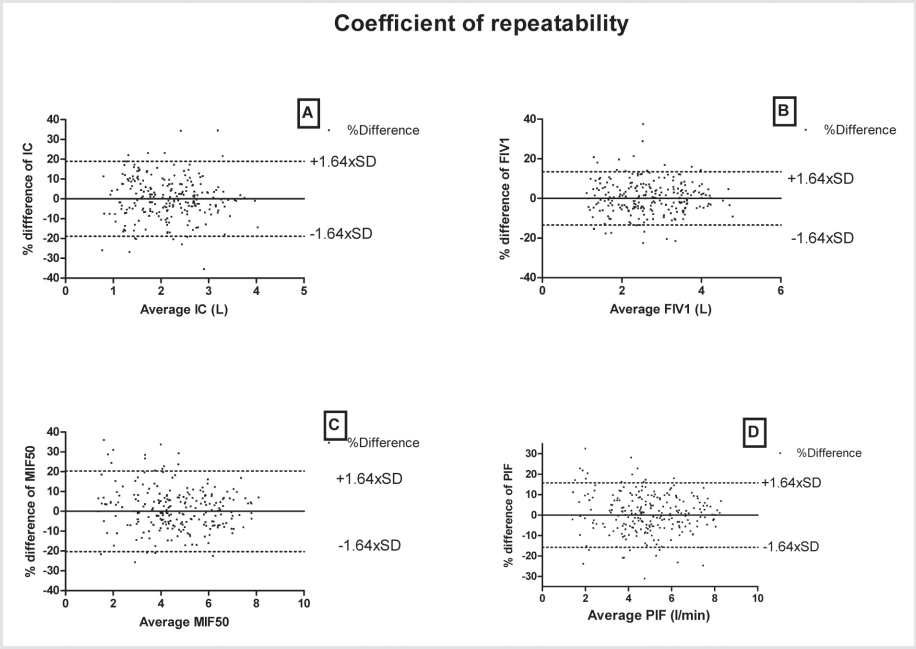


Figure 4. *Panels A–D show Bland–Altman plots of the percentage difference versus the average of the IC, FIV1, MIF50 and PIF. The coefficient of repeatability (CR) corresponds with the dotted lines; in panel A, e.g., the coefficient of repeatability takes the value = 18.9%, which is 1.64xSD.*
IC- Inspiratory Capacity; FIV1= Forced Inspiratory Flow in One second; MIF50= Maximal Inspiratory Flow at 50% ; PIF= peak Inspiratory Flow; SD=standard deviation.

In the same way as the one-hour coefficients of repeatability, the intra-day coefficients of repeatability and the in-between day coefficients of repeatability are investigated.

The intra-day random variabilities expressed as the coefficients of repeatability (CR) for the lung function parameters are: IC, 21%; FIV1, 18%; PIF, 18% or 0.90 l/; and MEF50, 21% (Table 4).

The inter-day random variabilities expressed as the coefficients of repeatability (CR) for the lung function parameters are: IC: 23%, FIV1: 14%, PIF: 18% and MEF50: 21% (Table 4).

Table 4. *Coefficients of repeatability (CR) retested after one hour (n=237), after five hours (n=79) and after 3-10 days (n-76) (in % from initial value/absolute value)*

Parameter	Retest after one hour	Retest after five hours	Retest another day
FEV1	11.8	13.7	17.9
IC	18.9	21.3	22.7
FIV1	13.5	17.9	18.0
PIF	17.9	17.8/ 0.90	17.9/ 0.85
MIF50	20.4	21.0	19.2

FEV1=Forced Expiratory Flow in one second; FIV1= Forced Inspiratory Flow in One second; IC- Inspiratory Capacity; MIF50= Maximal Inspiratory Flow at 50% PIF= peak Inspiratory Flow.

Discussion

Type of scatter

This study has shown that within the same subject, differences in lung function parameters (IC, FIV1, MIF50 and PIF) before and after one hour can more appropriately be described when taken as the percentages of the initial values than as differences in the absolute values because of the more homoscedastic scatter. All measured parameters showed a scatter that was significantly dependent on the average parameter value and thus is heteroscedastic if we present the differences in liters or L/sec. On the other hand, if we represent the difference as the percentage of the average value (which is nearby the initial value), we found no significant dependence on the parameter value for IC and less dependence on the values for the FIV1 and MIF50.

Several studies have addressed the variability of lung function parameters, especially on forced expiration [15-19], but used the variation coefficient instead of the method described by Bland and Altman [11]. Therefore, the type of spread was not investigated. In the ATS-ERS statement, the method of Bland and Altman is described as the preferred method for investigating

the random variation, and this method makes the type of spread visible [5,11].
The only exception is the PIF, which displays a slightly steeper slope when expressed as the percentage difference.

Random variability
The one-hour repeatability is by far the most important random variability because most interventions we are interested in, such as bronchodilator response, can be measured within one hour. Subjects must at least exceed this random variation before it can be decided that an improvement of an intervention can be attributed to that intervention. We did not find any CR for inspiratory parameters in the literature.
We decided to pool our patients' data (with one-hour differences measured at three time-points a day per patient) to obtain more data points and, thus, more reliable results. This pooling was possible because we found no significant differences between the group means and spreads and no significant differences of the parameters between the measurements between 9 a.m. and 2 p.m. This result is in contrast with Calverley et al. [20] and van Noord et al. [21], who found significantly lower values at 3 and 6 a.m.; however, we did not measure at these hours.
We chose the CR instead of the more popular variation coefficient because it more precisely reflects the repeatability and provides a graphical representation of the type of scatter, as stated by Bland and Altman [11] and the recommendations of the ERS-ATS committee [5].
The CR for the PIF is less than that for the MIF50 which may be because the MIF50 is situated near the PIF in maximal inspiratory flow volume curves but is seldom exactly aligned; thus, the MIF50 demonstrated more spread.
Whether PIF improvement is therefore more sensitive to bronchodilators than MIF is not answered by this study.
The intra-day coefficients of repeatability are important to know when we are performing interventions that take more than one hour, i.e., medications such as theophylline, tiotroprium or other interventions that take more time to retest.
We selected the 9-to-14-hour difference because all parameters as group means did not change during this interval. There was a small but significant decrease in some parameters (FEV1 and MIF50) on the 9-to-15-hour interval; therefore, we took the 9-to-14 as our difference. We think that this

decrease in some parameter values may be due to the fact that patients at the end become tired of repeating this procedure five times a day, during which time they were not allowed to take any bronchodilator drugs, or that there may be some circadian effect [20,21]. The higher intraday CR value, than the one hour CR value, could be expected because of the greater time interval.

Improvements of interventions taking more than one day can be considered as beyond random variation when the inter-day coefficients of repeatability are taken into account. Between the two days, patients did not change their medication and no exacerbations occurred. A weak point of these CR is that we were unable to see all patients on an exact inter-day interval of one week because we were dependent on when our patients were able to visit our outpatient department again. As the smallest interval was three days, and the greatest interval was eight days, the analyses were all conducted within two weeks. In general, the longer the interval between the two measurements (from one hour to several days), the greater the CR obtained. This result may be caused by the longer time period for weather to influence the patients or other effects of irritants in the environment.

Similar to the one-hour and intra-day random variation, we were unable to find the inter-day random variation on inspiratory parameters in the literature.

Limitations of this study

The subjects in this study include the investigation of only patients with COPD, so it does not extend to normal patients or those with asthma or restrictive disease. The random variation in these groups may be different.

The type of scatter was only examined after one hour, and it may be different when other intervals are taken into account.

The wash-out time for Tiotropium was 24 hours, although some investigators used 48 hours for this drug. We think this 24 hours time period had limited influence on the test-retest results.

We used all data including the outliers to construct the Bland and Altman plots; in small samples, this can influence the linear regression of the transformed Bland and Altman plots when the outliers are in the lower or upper zones of the average parameter value.

Conclusions

Differences in lung function parameters (IC, FIV1, MIF50 and PIF) are described with less dependence on the parameter values when taken as percentages from the initial values than as absolute difference values.
The random variation expressed as coefficients of repeatability for several time intervals are presented.

Reference List

1. Fabbri LM, Hurd SS: Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update5. *Eur Respir J* 2003, 22: 1-2.

2. Archibald CJ, Guidotti TL: Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab* 1987, 1: 45.

3. Guyatt GH, Townsend M, Nogradi S, Pugsley SO, Keller JL, Newhouse MT: Acute response to bronchodilator. An imperfect guide for bronchodilator therapy in chronic airflow limitation. *Arch Intern Med* 1988, 148: 1949-1952.

4. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162: 216-220.

5. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R *et al*.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26: 948-968.

6. Taube C, Kanniess F, Gronke L, Richter K, Mucke M, Paasch K *et al*.: Reproducibility of forced inspiratory and expiratory volumes after bronchodilation in patients with COPD or asthma. *Respir Med* 2003, 97: 568-577.

7. O'Donnell DE, Lam M, Webb KA: Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998, 158: 1557-1565.

8. O'Donnell DE, Lam M, Webb KA: Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999, 160: 542-549.

9. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A *et al*.: Standardisation of spirometry. *Eur Respir J* 2005, 26: 319-338.

10. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Recommendations for the measurement of FIV(1) values in chronic obstructive pulmonary disease. *Respiration* 2008, 76: 46-52.

11. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1: 307-310.

12. Celli BR, MacNee W: Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004, 23: 932-946.

13. Quanjer Ph.H., Tammeling G.J.: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993, 6: 5-40.

14. Hadcroft J, Calverley PM: Alternative methods for assessing bronchodilator reversibility in chronic obstructive pulmonary disease. *Thorax* 2001, 56: 713-720.

15. Anthonisen NR, Wright EC: Bronchodilator response in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1986, 133: 814-819.

16. Dompeling E, van Schayck CP, Molema J, Akkermans R, Folgering H, van Grunsven PM *et al*.: A comparison of six different ways of expressing the bronchodilating response in asthma and COPD; reproducibility and dependence of prebronchodilator FEV1. *Eur Respir J* 1992, 5: 975-981.

17. Enright PL, Connett JE, Kanner RE, Johnson LR, Lee WW: Spirometry in the Lung Health Study: II. Determinants of short-term intraindividual variability. *Am J Respir Crit Care Med* 1995, 151: 406-411.

18. Nickerson BG, Lemen RJ, Gerdes CB, Wegmann MJ, Robertson G: Within-subject variability and per cent change for significance of spirometry in normal subjects and in patients with cystic fibrosis. *Am Rev Respir Dis* 1980, 122: 859-866.

19. Sourk RL, Nugent KM: Bronchodilator testing: confidence intervals derived from placebo inhalations. *Am Rev Respir Dis* 1983, 128: 153-157.

20. Calverley PM, Lee A, Towse L, van Noord J, Witek TJ, Kelsen S: Effect of tiotropium bromide on circadian variation in airflow limitation in chronic obstructive pulmonary disease. *Thorax* 2003, 58: 855-860.

21. van Noord JA, Aumann JL, Janssens E, Verhaert J, Smeets JJ, Mueller A *et al*.: Effects of tiotropium with and without formoterol on airflow obstruction and resting hyperinflation in patients with COPD. *Chest* 2006, 129: 509-517.

5

Pursed-lip breathing improves inspiratory capacity in COPD

Frank J. Visser, Sunil Ramlal, P.N. Richard Dekhuijzen, Yvonne F. Heijdra

Respiration 2011; 81: 372-378.
PMID: 20639621

Abstract

Background: In patients with severe chronic obstructive pulmonary disease (COPD), pursed-lips breathing (PLB) improves the pulmonary gas exchange and hyperinflation measured by electro-optic coupling. The response to PLB in inspiratory lung function tests is not known. **Objectives:** The purpose of this study was to measure the effect of PLB on inspiratory parameters.

Methods: Thirty-five subjects with stable COPD and a forced expiratory volume in first second (FEV_1) <50% of the predicted value were tested for the following primary parameters before and immediately after PLB, and 5 min later: forced inspiratory vital capacity, inspiratory capacity (IC), forced inspiratory volume in first second, maximal inspiratory flow at 50% of VC, and peak inspiratory flow. Patients were also tested for the following secondary parameters: vital capacity (VC), FEV_1, breathing frequency (BF), end-tidal CO_2 tension (ET-CO_2), and oxygen saturation (SO_2).

Results: Of all the primary parameters only IC ($p = 0.006$) improved significantly; with regard to the secondary parameters, the mean SO_2 was improved by 1% ($p = 0.005$) and the mean ET-CO_2 and BF decreased significantly ($p < 0.0001$ for both) to 3.2 mm Hg and 3.1 breaths/min, respectively. After 5 min the effects diminished.

Conclusion: Improved IC after PLB indicate less hyperinflation in patients with severe COPD; there was no effect on parameters of flow.

Introduction

Chronic obstructive pulmonary disease (COPD) is a preventable and treatable disease but is still the fourth leading cause of death in the world [1]. The severity of COPD is defined by the degree of expiratory airflow limitation. Airflow obstruction is essential for diagnosis, and forced expiratory volume in first second (FEV1) provides a useful description of the severity of the pathological changes in COPD [1]. However, FEV1 is not very well correlated with changes in dyspnea. Inspiratory parameters may be more sensitive in relation to dyspnea as was published by Taube et al. [2]. They are also sensitive to bronchodilators in patients with COPD [2-4].

Pursed-lips breathing (PLB) is a breathing exercise and an item of patient education in rehabilitation programs [5, 6]. PLB may improve pulmonary gas exchange [7, 8] and reduce the breathing frequency (BF) and end-expiratory volume measured by optoelectronic plethysmography, thereby decreasing hyperinflation [9, 10]. A decrease in dyspnea and an increase in tidal volume are other consequences of PLB in patients with moderate-to-severe COPD [6]. Additionally, a faster recovery from dyspnea and a slower respiratory rate were found after walking with PLB [11].

The physiologic changes induced by PLB cause an increased intrabronchial pressure during expiration and, as a consequence, may increase the bronchial diameter and thus improve the inspiratory and expiratory flow. The positive intrabronchial pressure prevents the collapse of the bronchi upon expiration and may therefore decrease the closing volume and improve the inspiratory capacity (IC) and vital capacity (VC).

We do not know how long this effect remains after PLB; however, we think it is maintained for approximately 5 min during quiet breathing (except when the patient performs a forced expiration). We wonder whether FEV1 changes at all during PLB because of the compression due to a negative intrabronchial pressure causing airway collapse. We hypothesized that inspiratory parameters could be improved by PLB resulting in a decrease in dyspnea.

The aim of this study was to evaluate the effect of PLB in patients with severe-to-very-severe COPD (GOLD stages 3 and 4) on the following inspiratory parameters: forced inspiratory volume in first second (FIV1), IC, maximal inspiratory flow at 50% of VC (MIF50), and peak inspiratory flow

(PIF); secondary outcome parameters included FEV1, forced VC (FVC), oxygen saturation, end-tidal CO2 tension (ET-CO2), BF, and dyspnea.

Methods

A total of 35 consecutive patients who met the GOLD criteria for COPD were recruited from our outpatient clinic. Inclusion criteria were: GOLD stages 3 and 4, reversibility of FEV1 <12% of the predicted normal value and <200 ml, age ≥40 years, smoker or former smoker (≥10 pack-years), and stable disease. Patients on oral corticosteroids or antibiotics in the month before inclusion and patients with symptomatic heart failure, respiratory diseases other than COPD, a history of asthma, allergic rhinitis, and active cancer (except basal cell carcinoma of the skin) or with spontaneous PLB were excluded. The study was approved by the hospital's medical ethical committee and all patients gave their informed consent.

Study Design

Patients were asked not to use short-term bronchodilators 6 h prior to the study and long-term bronchodilators were stopped at least 12 h before the study. The use of tiotropium bromide and theophylline was not allowed 24 h prior to the spirometric test.

Patients were asked to rest and breathe quietly for at least 2 min before the start of the test, followed by the recording of basal values for BF, ET-CO2, and oxygen saturation (SO2) as well as lung function tests for the inspiratory and expiratory parameters FIV1, IC, forced inspiratory vital capacity (FIVC), PIF, MIF50, and FEV1.

After these measurements patients rested for 5 min and thereafter they learned the PLB procedure with the following instructions: 'Sit straight and relax your neck and shoulders. Lean with your arms on the arm rests of your chair. Breathe quietly in through your nose and out by means of pursed lips. During the inspiration your mouth should be closed. The expiration should be about 2 times longer in duration than the inspiration'.

After these instructions the patient was asked to demonstrate the PLB procedure; if the procedure was not adequately performed, the assistant corrected the patient by instructing him on what to do until the correct procedure was learned by the patient.

The values during PLB were recorded as follows: the patient practiced PLB for 2 min followed by 1 inspiratory maneuver to obtain the inspiratory

parameters; this process was repeated until 5 adequate inspiratory flow curves were obtained (fig. 1). The largest FIV1, IC, FIVC, PIF, and MIF50 were recorded. Responses to the visual analog scale (VAS), as well as SO2, ET-CO2, and BF were recorded just before the 5th inspiratory maneuver

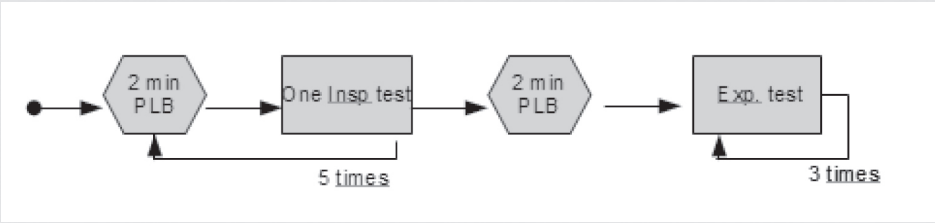


Figure 1. *Method of measurement of the lungfunction parameters. Two minutes of PLB is followed by one forced inspiratory maneuver this is repeated until 5 inspiratory maneuvers are obtained, 3 forced expiratory maneuvers are obtained after another 2 min. of PLB. Insp.=inspiratory min=minutes, Exp.=expiratory.*

Thereafter, the patient practiced PLB for 2 min followed by the expiratory maneuver to obtain the expiratory parameters. This process was repeated until 3 adequate expiratory flow curves were obtained. The largest FEV1 and FVC were recorded. Five minutes after the last measurement the same parameters were recorded along with the responses to the VAS in order to obtain the post-PLB values.

Pulmonary Function Tests

Lung function was measured both at forced expiration and inspiration as follows: a 3-liter calibration syringe was used at 3 different emptying and filling speeds to check linearity, as recommended by American Thoracic Society (ATS) and European Respiratory Society (ERS) standards. The ambient (room) temperature was measured before each test session to allow body temperature, pressure, and saturation corrections to be applied to the flows and volumes.

To measure the basal and post-PLB values of FVC and FEV1, patients performed as many maneuvers as needed (with a maximum of 8) to achieve 3 adequate and acceptable flow-volume curves, according to conventional ATS/ERS criteria.

For inspiratory parameters, 5 maximal forced inspirations after a slow and maximal expiration were obtained. Maximal inspiration was obtained when a plateau was reached or after at least 8 s of inspiration.

In order to obtain proper inspiratory parameters after a slow expiration, we started the measurement during slow expiration and stopped the procedure when the patient reached FIVC, as otherwise the software of the V-MAX20 spirometer (SensorMedics, ViaSys, Conshohocken, Pa., USA) would reject the obtained values.
If during the inspiratory maneuvrs VC was reached before FIV1, then FIV1 = VC. The largest FVC, FEV1, and FIV1 were recorded. For the predicted FEV1 and FVC, the normal values of the European Community for Steel and Coal were used [12].
The flow-volume curves were measured with a V-MAX20 spirometer (Sensor Medics). ET-CO2 was recorded with a Nellcor N1000 oximeter (Nellcor Puritan Bennett, Inc., Pleasanton, Calif., USA). SO2 was recorded with a Nellcor NPB40 pulse oximeter (Nellcor Puritan Bennett)
Visual Analog Scale
The patients were asked to fill out a VAS [2, 13]. On the 10-cm long VAS scale the middle represents no change, and the left and right edges of the line represent the most dyspnea and least dyspnea, respectively.
Statistics
The differences between inspiratory and expiratory parameter values before and after PLB were calculated with a 2-tailed paired Student t test. $p < 0.05$ was defined as a statistically significant difference. Correlations with the VAS scale (Spearman's rank correlation test) were determined. VAS scores are presented as means and confidence intervals (CI) of the means. The D'Agostino-Pearson omnibus normality test was used to check whether the distribution of the VAS scores was normal. For statistical calculations we used GraphPad Prism5 for Windows (www.graphpad.com).

Results

Of the 35 patients in the study, 2 were not able to learn the PLB procedure and 1 was not able to perform the inspiratory lung function maneuver. Therefore, 32 patients were eligible for analysis. Twenty-five patients had GOLD stage 3 and 7 had GOLD stage 4 COPD. The clinical and demographic characteristics are summarized in table 1.

Table 1 *Clinical and demographic characteristics*

	Mean (SD)
Subjects	32
Female/ Males	8/24
Age yrs	63.9 (7.5)
Height cm	170.8 (9.7)
Weight kg	75.9 (15.8)
FEV1 L	1.08 (0.37)
FEV1 % pred	37.15 (11.6)

Change in Inspiratory Parameters during PLB and 5 min Later.

During PLB, we found a significant improvement in IC with a mean increase of 89 ml (range --190 to +570); 6 patients had an increase of 200 ml or more. MIF50 showed a significant mean decrease of 170 ml/min. The other parameters were not significantly altered by PLB.

Five minutes later, none of the inspiratory parameters showed any significant improvement in relation to the basal values (before PLB). The IC was still 61 ml higher than at baseline but a 2-tailed paired t test showed that this difference lacked significance (p = 0.061). When we compared the changes in parameters during PLB and 5 min after PLB, we found a mean change in IC of 28 ml (p = 0.237, not significant). The results are summarized in table 2.

Table 2

	Differences in inspiratory parameters before and during Pursed Lip Breathing			Differences in inspiratory parameters before and 5 minutes after Pursed Lip Breathing		
Parameter	**Mean change of differences**	**95% Confidence Interval**	**P value**	**Mean change of differences**	**95% Confidence Interval**	**P value**
FIV1 L	0.061	-0.0041 to 0.073	0.57 ns	0.025	-0.044 to 0.094	0.47 ns
IC L	0.089	0.038 to 0.140	0.006**	0.061	-0.004 to 0.126	0.061 ns
MIF50 L/sec	-0.17	0.0 to 0.34	0.049 *	-0.15	-0.371 to 0.070	0.087 ns
PIF L/sec	-0.084	-0.26 to 0.09	0.34 ns	-0.007	-0.220 to 0.21	0.946 ns

*P vaulue < 0.001 (extremely significant) is summarised with ***; P value 0.001 to 0.01 (very significant) summerised with ** and 0.01 to 0.05 (significant) summarised with **

Change in Secondary Parameters during and 5 min after PLB.
Expiratory lung function parameters FEV1 and FVC did not show significant differences (mean changes in differences -11 and +59 ml, respectively). However, SO2, end-tidal pCO2, and BF all showed small but significant improvements during PLB.
Five minutes after PLB the improvements diminished somewhat, except in FVC which showed a significant improvement compared to the basal value (mean change in differences 105 ml; p = 0.009); however, there was no significant improvement compared to the value immediately after PLB (mean change in differences 46 ml; p = 0.143). The results are summarized in table 3.

Table 3

	Differences in secondary parameters before and during Pursed Lip Breathing			Differences in secondary parameters before and 5 minutes after Pursed Lip Breathing		
Parameter	**Mean change of differences**	**95% Confidence Interval**	**P value**	Mean change of differences	95% Confidence Interval	P value
FEV1 L	-0.011	-0.056 to 0.024	0.52 ns	-0.016	-0.052 to 0.019	0.341 ns
FVC L	0.059	-0.029 to 0.148	0.182 ns	0.105	0.0285 to 0.183	0.0093 **
Oxygen saturation %	0.97	0.463 to 1.474	0.0005 ***	0.375	-0.173 to 0.923	0.172 ns
End-tidal CO_2 mmHg	-3.281	-4.298 to -2.26	<0.0001 ***	-2.59	-3.60 to -1.59	<0.0001 ***
Breathing frequency b/min	-3.063	-4.066 to -2.059	<0.0001 ***	-1.31	-2.421 to 0.204	0.022 *

*P vaulue < 0.001 (extremely significant) is summarised with ***; P value 0.001 to 0.01 (very significant) summerised with ** and 0.01 to 0.05 (significant) summarised with **

Correlations between the Dyspnea Score and Improving Parameters with PLB.
We correlated those parameters that showed significant improvements during or after PLB with the subjective change in the patients' feelings of dyspnea. None of the parameters showed a significant correlation with the patients' feelings of dyspnea. Only SO2 showed a tendency to correlate weakly with the VAS score (--0.038, p = 0.08, not significant). The results are summarized in table 4.

Table 4 *Correlations between dyspnea (VAS) and change in parameter values*

	Immediately after PLB		**5 minutes after PLB**	
Parameter	Correlation	P value	Correlation	P value
FVC	0.249	0.17	0.146	0.42
IC	0.061	0.74	0.299	0.096
Oxygen saturation %	0.377	0.059	0.314	0.08
End-tidal CO_2 mmHg	-0.201	0.27	-0.108	0.554
Breathing frequency b/min	-0.044	0.81	0.104	0.571

VAS Score Immediately after PLB and 5 min after PLB.
Patients recorded an improvement on the VAS scale immediately after PLB with a mean of 7.8 mm (CI 3.3-12.2) out of 50 mm. Patients recorded an improvement on the VAS scale 5 min after PLB with a mean of 7.6 mm (CI 2.6-12.5) out of 50 mm. The distribution of the VAS scores was normal.

Discussion

Change in Inspiratory Parameters Immediately after PLB

We found a significant improvement in IC following PLB. However, 9 patients showed a decrease in IC ranging from 10 to 190 ml; 4 of them had a decrease of more than 100 ml.

To our knowledge, no data exist on inspiratory pulmonary function tests after PLB in the literature, but of all the inspiratory parameters measured (FIV1, IC, MIF50 and PIF) only IC showed improvement. IC is a static lung function parameter that is also a marker of hyperinflation. Optoelectronic plethysmography (OEP) showed a significant reduction (mean ± SD) in the end-expiratory volume of the chest wall [9] during PLB (-0.33 ± 0.24 liters; $p < 0.000004$). This finding of a lower end-expiratory volume by OEP is supported by our finding of an increase in IC after PLB. However the change in OEP volume of 0.33 liters was higher than our mean change of 0.098 liters. This difference may be partly attributed to patient selection, as we only included GOLD stages 3 and 4, and to the other types of measurements that were performed (change in chest wall dimensions). Our study also supports reduced hyperinflation (improved IC) following PLB as reported

previously. This reduction in hyperinflation may also be responsible for the improved oxygen cost of breathing [14] and the faster recovery after walking [15].
MIF50 showed a significant mean decrease of 170 ml/s ($p = 0.049$). This change in MIF50 was the opposite of what we expected. We speculate that this effect may be caused by reflex bronchoconstriction (a stretching of the J receptor in the bronchial wall caused by higher intrabronchial pressure during PLB).All other parameters of dynamic (forced) lung function were not significantly changed after PLB.

Change in Inspiratory Parameters 5 min after PLB
All inspiratory lung function parameters lacked significant improvement after 5 min in relation to the basal values (before PLB). The IC was still 61 ml higher than at baseline but a 2-tailed paired t test showed that this difference lacked significance ($p = 0.061$); when we compared this value to the IC during PLB we found no significant decrease either. Thus, after 5 min some of the initial improvement in IC had disappeared. We did not take any measurements later to see how long the improvement due to PLB lasted, nor did we find any clues in the literature to answer that question.

Change in Secondary Parameters Immediately after PLB
Expiratory lung function parameters like FEV1 and FVC did not show change after PLB. No studies are available on this subject; therefore, we are not able to compare our results with those of others.
In contrast to FEV1 and FVC, SO2, end-tidal pCO2, and BF showed small but significant improvements similar to those reported previously [16]. The reason for a better SO2 may be a lower cost of oxygen due to less hyperinflation (less work breathing). Less hyperinflation improves lung compliance which may be the explanation for the decrease in BF.

Change in Secondary Parameters 5 min after PLB
Five minutes after PLB the improvements in BF, end-tidal pCO2, and SO2 diminished again. FVC, however, improved compared to the basal value. This improvement in FVC surprised us, but might also reflect less hyperinflation. Why this change reached significance after 5 min and not immediately after PLB is a question we cannot answer. As stated earlier, we did not find any data on these parameters in the literature.

Dyspnea and Correlations with Changes in Parameter Values.

We found a mean difference of 4.85 ml in the VAS score after test-retest, which was significantly less than the 7.75ml mean difference found immediately after PLB (Mann-Whitney test; $p < 0.001$). This is compatible to the improvement in dyspnea that we found in the literature [6, 16]. However, no significant correlations were found between significantly changed parameters after PLB and the feeling of change in dyspnea recorded by the patients. Furthermore, as this study was underpowered to find significant changes, very weak correlations remained; thus, we do not know which of the parameters contributed to a reduced feeling of dyspnea in patients after PLB. This issue was reviewed by Dechman and Wilson [16] and they found only 1 article, by Ingram et al. [17], which suggested that the higher collapsibility of bronchial airways in responders compared with nonresponders could be responsible.

Spahija et al. [18] found a strong and significant association between change in the end-expiratory lung volume and the VAS score during exercise. However, they only tested 8 patients with COPD, 6 of whom had an FEV1 <50% of the predicted value. From their data, we calculated the association between the change in IC and the change in the VAS score and we also found a significant association; however, when we omitted patient 7 of their data (because this case had an outlying change in the dyspnea VAS score compared with the other 7 subjects), the significant association vanished. We think that much more data of patients with severe COPD are needed in order to obtain robust results. The VAS dyspnea did not change before or after PLB at rest (which is in line with what we found). Their method of VAS measurement was different; they used an absolute VAS scale (0-10) 2 times and we used a VAS scale 1 time to express the difference (less or more dyspnea).

Bianchi et al. [10] asked 30 patients with stable COPD to perform PLB at rest and found, in 19 patients, a reduction in the end-expiratory volume of the chest wall (VeeCW), corresponding with less hyperinflation after PLB and an increase in tidal volume. Overall, he found an association between a decrease in VeeCW and a change in the BORG scale. We did not find this association, but we used a VAS scale and not a BORG scale, and while we measured IC via spirometry they used OEP. Patients who hyperinflated during PLB had better FEV1 as a percentage of the predicted values (FEV1 %pred) in their basal values. In the 4 patients with a decrease in IC of more

than 100 ml during PLB we found a mean FEV1 %pred of 42% compared with 37% in the whole group; we also found a mean change of 8.2 mm in their VAS dyspnea scores (range 0--20) compared with 7.8 mm in the whole group.
In another study Bianchi et al. [9] analyzed 22 patients with COPD and found that the patients with a greater reduction in hyperinflation were the patients with more severe airway obstruction. He also described a longer breathing cycle after PLB (hence a lower BF). We also found a lower BF after PLB but this was not associated with less dyspnea in our sample of COPD patients.
In our study 10 patients with a decrease of 5 or more breaths/min in their BF had a mean change in VAS score of 9.3 mm compared with 7.8 mm for the whole sample. Bianchi et al. used OEP and a Borg scale in their study. Despite the fact that we were unable to associate less dyspnea with less hyperinflation at rest, Bianchi et al. found this association in their sample of patients and Spahija et al. found an association between PLB and less dyspnea during exercise but not at rest. We think a larger sample of severe COPD patients is needed to clarify the association between PLB and the decrease in dyspnea.

Conclusions

This study showed that there was an improvement in IC after PLB, supporting the idea of a decreased hyperinflation in patients with severe COPD and a possibly higher collapsibility of the bronchial airways. SO2, end-tidal pCO2, and BF also improved. We were not able to correlate these changes with a decreased VAS dyspnea score, however.

Reference List

1. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2007. Available from: http://www.goldcopd.org.

2. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162: 216-220.

3. O'Donnell DE, Lam M, Webb KA: Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999, 160: 542-549.

4. O'Donnell DE, Lam M, Webb KA: Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998, 158: 1557-1565.

5. Nici L, Donner C, Wouters E, Zuwallack R, Ambrosino N, Bourbeau J *et al*.: American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation.

   Am J Respir Crit Care Med 2006, 173: 1390-1413.
6. Troosters T, Casaburi R, Gosselink R, Decramer M: Pulmonary rehabilitation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005, 172: 19-38.

7. Breslin EH: The pattern of respiratory muscle recruitment during pursed-lip breathing. *Chest* 1992, 101: 75-78.

8. Fagevik OM, Westerdahl E: Positive expiratory pressure in patients with chronic obstructive pulmonary disease--a systematic review. *Respiration* 2009, 77: 110-118.

9. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Grazzini M *et al*.: Chest wall kinematics and breathlessness during pursed-lip breathing in patients with COPD. *Chest* 2004, 125: 459-465.

10. Bianchi R, Gigliotti F, Romagnoli I, Lanini B, Castellani C, Binazzi B *et al*.: Patterns of chest wall kinematics during volitional pursed-lip breathing in COPD at rest. *Respir Med* 2007, 101: 1412-1418.

11. Garrod R, Dallimore K, Cook J, Davies V, Quade K: An evaluation of the acute impact of pursed lips breathing on walking distance in nonspontaneous pursed lips breathing chronic obstructive pulmonary disease patients. *Chron Respir Dis* 2005, 2: 67-72.

12. Quanjer Ph.H., Tammeling G.J.: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993, 6: 5-40.

13. Noseda A, Schmerber J, Prigogine T, Yernault JC: Perceived effect on shortness of breath of an acute inhalation of saline or terbutaline: variability and sensitivity of a visual analogue scale in patients with asthma or COPD. *Eur Respir J* 1992, 5: 1043-1053.

14. Jones AY, Dean E, Chow CC: Comparison of the oxygen cost of breathing exercises and spontaneous breathing in patients with stable chronic obstructive pulmonary disease. *Phys Ther* 2003, 83: 424-431.

15. Gigliotti F, Romagnoli I, Scano G: Breathing retraining and exercise conditioning in patients with chronic obstructive pulmonary disease (COPD): a physiological approach. *Respir Med* 2003, 97: 197-204.

16. Dechman G, Wilson CR: Evidence underlying breathing retraining in people with stable chronic obstructive pulmonary disease. *Phys Ther* 2004, 84: 1189-1197.

17. Ingram RH, Jr., Schilder DP: Effect of pursed lips expiration on the pulmonary pressure-flow relationship in obstructive lung disease. *Am Rev Respir Dis* 1967, 96: 381-388.

18. Spahija J, de MM, Grassino A: Effects of imposed pursed-lips breathing on respiratory mechanics and dyspnea at rest and during exercise in COPD. *Chest* 2005, 128: 640-650.

6

Reversibility of inspiratory lung function parameters after short-term bronchodilators in COPD

S.K. Ramlal, F. J. Visser, W. C. Hop, P. N. R. Dekhuijzen, Y. F. Heijdra

Respir Physiol Neurobiol 2010; 173: 58-63.

Abstract

Background: The responsiveness of short-term bronchodilators on inspiratory lung function parameters (ILP), Forced Inspiratory Volume in one second (FIV_1), Inspiratory Capacity (IC), Forced Inspiratory Flow at 50% of the vital capacity (FIF_{50}), Peak Inspiratory Flow (PIF) and the relationship with dyspnoea in COPD subjects is sparsely examined. The aim of this study was to assess the effects of inhaled salbutamol 400 mcg, ipratropium 80 mcg and placebo on ILP and FEV_1, and their relation with Visual Analogue Scale (VAS).

Methods: 85 subjects with stable COPD participated in a cross-over, randomized, double-blind placebo-controlled study. Spirometry was performed before and after inhalation of salbutamol, ipratropium and placebo. Changes in dyspnoea were measured with VAS. Primary analysis was done using 63 participants with absent reversibility.

Results: all ILP and FEV_1 improved significantly on bronchodilators except FIF_{50} after Ipratropium. After both bronchodilators, percent changes from initial values in IC were higher than corresponding changes in FEV_1, but the differences were not significant (all $p>0.06$).
Mean VAS score after bronchodilators and placebo showed significant improvements but did not significantly correlate with changes in lung function parameters. For each lung function parameter patients were further classified as responder or non-responder according to the degree of the change from baseline. Response rates did not significantly differ between the various ILP. Also no significant differences were found between responders and non- responders regarding dyspnoea after bronchodilators. This applied to all ILP as well as to FEV_1.

Conclusions: In subjects with COPD, all ILP and FEV_1 showed, just like the VAS score, significant improvements after bronchodilators. However ILP were not more sensitive than FEV_1 for detecting responders on bronchodilators and for changes in the VAS score.

Introduction

In Chronic Obstructive Pulmonary Disease (COPD) patients chronic inflammatory reaction and structural changes in central and peripheral airways and lung parenchyma are assumed to result in loss of elasticity of the airways leading to expiratory peripheral airway collapse, expiratory airflow limitation and hyperinflation [1,2]. Airway obstruction is worldwide measured with Forced Expiratory Volume in one second (FEV_1) and vital capacity (VC) ratio < 5 th percentile of predicted [3]. COPD is defined by FEV_1/FVC <70% after the inhalation of short-acting bronchodilator [2]. In COPD patients FEV_1 has a poor relationship with dyspnoea and is used, as a diagnostic tool for evaluation of bronchodilator therapeutic response and for spirometric classification of COPD severity [4-8]. FEV_1 is often used because of its simplicity in performance, its known random variation and its well established reference values. There is a lack of correlation between the FEV_1 and dyspnoea score in general and a low correlation between changes in FEV_1 and changes in dyspnoea score after bronchodilation. O'Donnell et al., found a correlation between the change of the Inspiratory Capacity (IC) and dyspnoea after bronchodilator and during exercise [9-11] Taube et al. found in stable COPD at rest better correlations between the inspiratory lung function parameter Forced Inspiratory Volume in one second (FIV_1) and dyspnoea measured with Visual Analogue Scale (VAS) than with the expiratory lung function parameter FEV_1 and dyspnoea [12].

We studied before the method of measurement of the FIV_1 and random variation of inspiratory lung function parameters (ILP). Responders were defined as subjects who must exceed the random variation [13].

Our hypothesis was that in COPD patients after inhalation of short-acting bronchodilators we would find more responders with inspiratory lung function parameters FIV_1, IC, Forced Inspiratory Flow at 50% of the vital capacity (FIF_{50}) and Peak Inspiratory Flow (PIF) than with FEV_1. Additionally we aimed to confirm the findings of Taube et al, who found a strong correlation between percentage change in FIV_1 and improvement of dyspnoea after bronchodilators.

The aim of our study was to investigate if the effect of short-acting bronchodilators on ILP was higher compared with FEV_1, on the number of responders and to investigate changes of these parameters for their correlation with dyspnoea scores.

Methods

Patients

Eighty-five stable COPD patients of the pulmonary outpatient clinic of the Canisius Wilhelmina hospital in Nijmegen, the Netherlands, were enrolled in this study performed in 2007 and 2008. The subjects had mild to very severe COPD (based on post-bronchodilator FEV_1) according to the GOLD guidelines [1]. Inclusion criteria were stable COPD, age between 40 and 80 years, current or former smoker with more than 10 pack-years and absent reversibility (an increase in FEV_1 that is both greater than 200 ml and 12% above the pre-bronchodilator FEV_1, according to historical records less than one year old) after short acting bronchodilators [3]. Stable COPD was defined as absence of exacerbations, no changes in COPD medications in the last 8 weeks, no use of oral corticosteroids in the last 2 months and no use of antibiotics in the last month. Also excluded were patients on oxygen, patients with allergic rhinitis or asthma, heart disease, neuromuscular disorders, malignancy or inability to answer questionnaires.

Study design

A double-blind, randomized placebo-controlled cross-over study was conducted. The randomization was performed by means of a computer generated list. The same subjects participated on three different days of the study, performing spirometric tests with reversibility testing on salbutamol, ipratropium or placebo. In a random order they were measured within a period of 2 weeks, received 4 puffs of 100 mcg salbutamol, ipratropium bromide 4 puffs of 20 mcg or 4 puffs of a placebo aerosol all with an Aerochamber plus. On these days standardized spirometry tests were performed before and after 30 minutes of the inhalation.

Pulmonary Function Testing

All subjects were asked not to use short-term bronchodilators 6 hours prior to the study and long-term bronchodilators at least 12 hour before the study. Tiotropium and theophylline were not allowed to use 24 hours prior to spirometric test.

Expiratory spirometry measurements were performed using standard techniques according to ERS and ATS guidelines [14,15]. The largest Forced Vital Capacity (FVC) and FEV_1 were recorded.

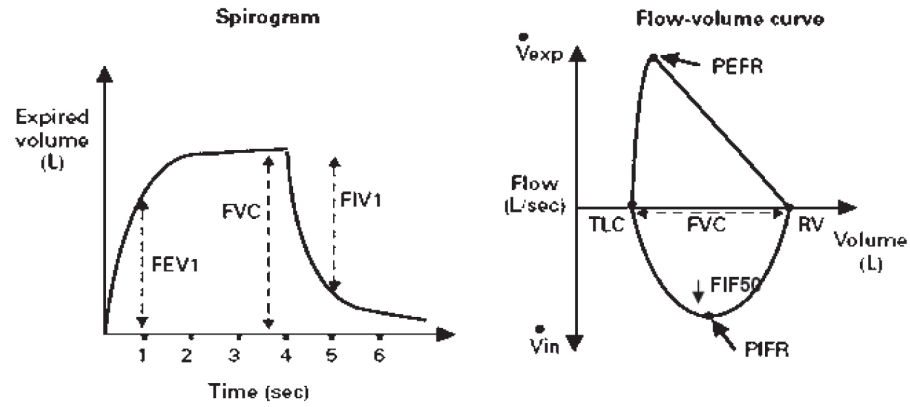


Figure 1 *shows a normal spirogram and a flow-volume curve.*

For the predicted FEV_1 and FVC the normal values of the European Respiratory Society were used (Quanjer and Tammeling, 1993). For ILP, maximal inspiratory flow was performed after slow and maximal expiration and at least five adequate manoeuvres were obtained (Visser et al., 2008) [13]. The best of the derived inspiratory parameters were further analyzed. Tests were performed with the patient in a seated position wearing a nose clip. Full inspiration was obtained when a plateau in the flow was reached or after at least 8 second duration of the inspiration. If during the inspiratory manoeuvres the VC was reached before FIV_1, then FIV_1 was considered to be equal to the FIVC. The largest FVC, FEV_1, Forced Inspiratory Vital Capacity (FIVC) and FIV_1 were recorded. For the predicted FEV_1 and FVC the normal values of the European Community for Steel and Coal were used [16]. IC was measured with the method described by Hadcroft et al. [17]. The flow-volume curves were measured with the V-MAX20 (Sensor Medics, ViaSys, and Conshohocken, PA, USA). For the three different measurements we used the same device and the measurements were performed by the same person at the same time of the day.

In order to obtain proper inspiratory parameters after a slow expiration we started the measurement during slow expiration and stopped the procedure when the patient reached FIVC, otherwise the software of the V-MAX20 rejects the obtained values [13]. The differences in FEV_1 and the ILP values (FIV_1 or FIVC, IC, FIF_{50} and PIF) after bronchodilatation were expressed both as changes in litres and as percent changes from initial value.

We defined a significant improvement within a subject (responder) as an

improvement of more than the Coefficient of Repeatability (CR). The one hour CR of the lung function parameters from initial value were for FEV_1 12 %, FIV_1 14 %, IC 19%, FIF_{50} 21 % and PIF 18 % [18]. The CR within one hour, intraday and interday, was determined by the method described by Bland and Altman [19]. The CR was determined as 1.64 times the standard deviation of the differences represented as absolute values or as percentage of the average values. We performed this one-tailed test instead of the two tailed test described because interventions we are interested in, should improve that parameter. Therefore we took 1.64 times SD instead of 1.96 times SD. This is analogue the way in which lower limits of normal are calculated[16]. CR was calculated instead of the more commonly used Coefficient of Variation (CV) because CV does not take into account the type of scatter that can be visualized by the scatter plots of Bland and Altman.

Prior to the test all subjects were instructed on the use of the VAS. Dyspnoea score with a VAS was taken 30 minutes after bronchodilator administration [20]. The VAS line is a 10 cm long horizontal line and analyzed ranges from -5 to + 5 cm. On the left side the label represents VAS= -5 (very much improved), in the middle representing VAS=0 (no change) and at the right end VAS = +5 (much worsened).

The medical ethical commission gave permission for this study and all patients gave written informed consent.

Statistical Analysis

All results are presented as mean value ± SD or indicated otherwise. Student's t- test for paired data was used to compare the changes in parameters after bronchodilators or placebo.

The one-sample t-test was used to compare mean changes with zero. Correlations between the change in VAS score and changes of lung function parameters after administration of bronchodilator and placebo were assessed with Spearman's rho test. SPSS for Windows, version 15, was used for analysis and p-values < 0.05 (two-sided) were considered statistically significant. Mann-Whitney test was used to test the differences between responders and non-responders regarding distribution of VAS for ILP after bronchodilators.

All data were primarily analysed in the group of patients who were found to be non-reversible. An analysis for the whole group of included patients was also done.

Results

Eighty-five COPD patients consented to participate in our study and had absent reversibility according to our patient records. Eighty-two of these patients completed the study. After completion of the study and unblinding the study medication at the analysis phase we discovered that a substantial fraction of the patients had a larger FEV_1 responsiveness than expected. It turned out that 19 patients had an FEV_1 increase ≥ 10% of the predicted value after the use of salbutamol and/ or after ipratropium.
For the primary analysis these 19 patients were excluded and the remaining 63 patients were analyzed. Patient's demographics and baseline pulmonary function of the 63 patients are presented in Table 1. These patients represent all four COPD GOLD classes.

Table 1. *Patient characteristics and baseline lung function parameters in 63 subjects*

Gender, Male/Female ratio (percentage)	48/15 (76% / 24%)
Age, years	66±9
Body Mass Index, kg/m²	27±6
Height, cm	170 ± 10
Weight, kg	77± 5
Smoker ex-/current, numbers (percentage)	27/36 (43% / 57%)
GOLD 1 patients, n	7
GOLD 2 patients, n	26
GOLD 3 patients, n	21
GOLD 4 patients, n	9
FEV_1, L	1.49±0.73
FEV_1 % predicted	51±20
FIV_1, L	2.68 ± 0.86
IC, L	2.18 ± 0.72
FIF_{50}, L/s	4.63 ± 1.60
PIF, L/s	4.96 ± 1.64

*Definition of abbreviations:*FEV1 = forced expiratory volume in 1 s, FIV1 = forced inspiratory volume in 1 s, IC = inspiratory capacity, FIF50 = forced inspiratory flow at 50% of the vital capacity, PIF= peak inspiratory flow. Values are given as mean ± SD. Baseline values were calculated as mean values of measured values at the three study days. GOLD = Global Initiative for Chronic Obstructive Lung Disease

Changes in litres and percent changes from initial values in lung function parameters after bronchodilators

Changes in litres (or L/s) and percent changes from initial values in lung function parameters and VAS scores after treatments are presented in Table 2. For salbutamol the range of changes (minimum-maximum) in FIV_1, FIF_{50} and PIF with FEV_1 was between 26 ml and approximately 229 ml. Ipratropium showed similar results with changes in litres ranging from 30 ml to 129 ml. From placebo the range of changes in ILP with FEV_1 was between 19 ml and 112 ml. All increases in litres (or L/s) after bronchodilators were significant ($p < 0.01$), except for changes in FIF_{50} after Ipratropium which was not significant ($p = 0.12$). Changes in litres or L/s after placebo were not significant except small decreases for IC ($p = 0.03$) and FIF_{50} ($p = 0.047$).

The percent changes from initial value in FEV_1 after bronchodilators and changes in the inspiratory parameters were all significantly (all $p<0.008$) different from zero except for FIF_{50} after ipratropium ($p= 0.06$). Percent changes after placebo for all lung function parameters were not significant (all $p > 0.16$).

Mean VAS score after both bronchodilators and placebo showed significant improvements (all $p < 0.001$) but did not significantly differ between each other (all $p>0.09$).

No statistically significant correlations were found between GOLD stages and changes in litres (or L/s) or relative changes of the ILP after both bronchodilators and placebo.

Relationships between changes in lung function parameters and dyspnoea (VAS scores) after bronchodilators or placebo

The relationships between VAS score and percent changes from initial values in FEV_1, FIV_1, IC, FIF_{50} and PIF after 30 minutes salbutamol and ipratropium is presented in Table 3.

Table 2. *Changes in litres (or L/s) and percent changes from initial value as compared with baseline values at 30minutes after salbutamol, ipratropium or placebo and visual analogue scale (VAS) for dyspnoea, in 63 subjects with stable COPD. Values are expressed as mean ± SEM.*

	changes in litres (or L/s)			**percent changes from initial value**		
	salbutamol	ipratropium	placebo	salbutamol	ipratropium	placebo
FEV_1	0.109±0.013	0.072±0.015	-0.003±0.012*	8±1	5±1	0±1*
FIV_1	0.135±0.031	0.102±0.027	-0.022±0.027*	6±1	4±1	0±1*
IC	0.170±0.030	0.132±0.029	-0.064±0.029	10±2	6±2	-2±1*
FIF_{50}	0.284±0.072	0.121±0.077*	-0.115±0.057	8±2	4±2*	-1±1*
PIF	0.338±0.068	0.201±0.075	-0.076±0.058*	8±2	4±2	0±1*

	salbutamol	ipratropium	placebo
mean VAS scores			
	-0.8±0.2	-0.9±0.2	-0.6±0.2

Definition of abbreviations:
FEV1 = forced expiratory volume in 1 s, FIV1 = forced inspiratory volume in 1 s, IC = inspiratory capacity (L),
FIF50 = forced inspiratory flow at 50% of the vital capacity (L/s), PIF = peak inspiratory flow (L/s). Changes in FEV1, FIV1, and IC are given in litres and FIF50 and PIF in L/s.
For all lung function parameters percent changes from initial values were calculated as (post bronchodilator -pre bronchodilator/pre bronchodilator)*100%.
*= changes after ipratropium or placebo which are not significant.
VAS= Visual analogue scale is given in mean ± SEM. VAS scale ranges from -5 (very much improved dyspnoea) to VAS= +5 (much worsened dyspnoea) and VAS=0 represents no change.

Table 3. *Correlations of VAS score with lung function parameters changes after salbutamol and after ipratropium in 63 patients. Data shown are Spearman correlations with p-value*

Δ	FEV_1 %	Δ FIV_1 %	ΔIC %	Δ FIF_{50} %	ΔPIF%
VAS score after salbutamol	-0.09	-0.003	-0.18	-0.006	0.06
p-value	0.47	0.98	0.15	0.96	0.65
VAS score after ipratropium	0.07	-0.08	0.23	-0.10	0.03
p-value	0.60	0.53	0.07	0.43	0.85

Definition of abbreviations:
Δ% represents = percent change of initial value
VAS= Visual analogue scale
FEV1 = forced expiratory volume in 1 s
FIV1 = forced inspiratory volume in 1 s
IC = inspiratory capacity (L)
FIF50 = forced inspiratory flow at 50% of the vital capacity (L/s)
PIF = peak inspiratory flow (L/s)

No statistically significant correlations were found between VAS scores and the percent changes from initial value after bronchodilators for all lung function parameters, neither for salbutamol nor for ipratropium.

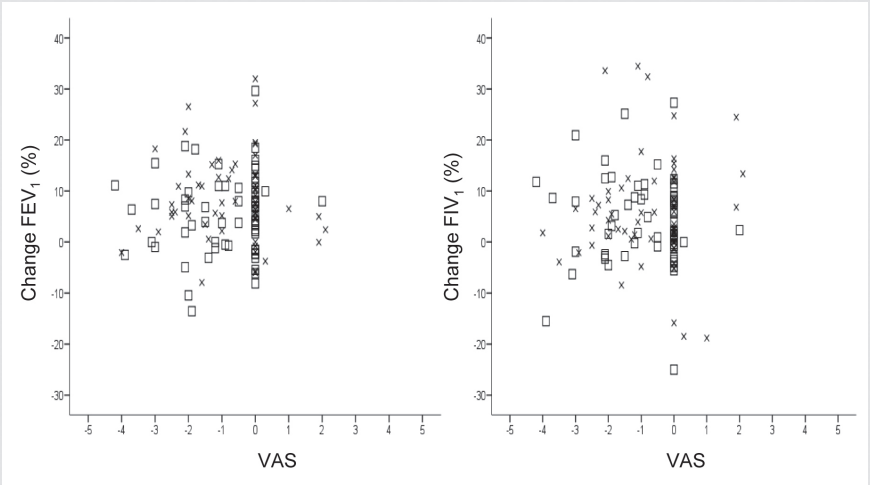


Figure. 2 *shows individual data for FEV_1 and FIV_1 changes versus VAS scores.*

For each lung function parameter (FEV_1, FIV_1, IC, FIF_{50} and PIF) patients were classified into responders or non- responders. Responders were defined as subjects with a change greater than the one hour CR from initial value after bronchodilators.

Using the Mann-Whitney test, no significant differences were found between responders and non- responders regarding the distributions of dyspnoea VAS scores for any of the lung function parameters, neither for salbutamol nor for ipratropium. These data are presented in Table 4. Neither for ipratropium nor for salbutamol we found significant differences between the various lung function parameters with regard to response rates.

Table 4. *VAS scores according to whether or not the change from baseline was greater than the one hour coefficient repeatability of the various lung function parameters in 63 patients*

	ipratropium 80 mcg		**salbutamol 400 mcg**	
response criterium	non-responder/ responder (numbers of patients)	Vas score non-responder/ responder	Non-responder/ responder (numbers of patients)	Vas score non-responder/ responder
FEV_1 response > 12 %*	53 vs 10 (84% vs 16%)	-0.8±1.2 vs -0.9±1.3	45 vs 18 (71% vs 29%)	-0.8±1.3 vs -0.8±0.9
FIV_1 response > 14 %*	58 vs 5 (92% vs 8%)	-0.8±1.2 vs -1.4±1.2	54 vs 9 (86% vs 14%)	-0.8±1.3 vs -0.3±1.1
IC response > 19%*	57 vs 6 (90% vs 10%)	-0.9±1.3 vs -0.9±1.2	50 vs 13 (79% vs 21%)	-0.7±1.2 vs -1.2±1.2
FIF_{50} response > 21%*	56 vs 7 (89% vs 11%)	-0.8±1.2 vs -1.3±1.1	56 vs 7 (89% vs 11%)	-0.8±1.2 vs -0.7±1.6
PIF response > 18 %*	43 vs 8 (84% vs 16%)	-0.9±1.3 vs -0.6 ±1.3	51 vs 12 (81% vs 19%)	-0.8±1.2 vs -0.6±1.5

Definition of abbreviations:
VAS= Visual analogue scale. VAS dyspnoea scores are given as mean ± SD. VAS ranges from -5 to + 5. VAS =0 represents no change in dyspnoea, VAS<0= less dyspnoea and VAS >0= more dyspnoea.
*= one hour coefficient of repeatability of the particular lung function parameter
FEV1 = forced expiratory volume in 1 s
FIV1 = forced inspiratory volume in 1 s
IC = inspiratory capacity (L)
FIF50 = forced inspiratory flow at 50% of the vital capacity (L/s)
PIF = peak inspiratory flow (L/s)

Essentially the same conclusions held when all analyses were repeated for the whole group of eighty-five included patients. The same applied to the subgroup of 51 patients of out of the whole group who met the stronger ATS/ERS criteria about non-reversibility [3] .

Discussion

This study showed that all ILP as well as FEV_1 have statistical significant improvements after short-acting bronchodilators in subjects with a stable COPD, except for FIF_{50} after ipratropium (Table 2).
For both bronchodilators the percent changes from initial value in IC were different corresponding changes in FEV_1. For all the ILP, we found a small amount of responders after both bronchodilators and the response rates for both bronchodilators did not differ significantly. Response rates for FEV_1 after both bronchodilators were also different corresponding response rates for ILP (Table 4).
The criteria of reversibility after bronchodilator in COPD are very confusing and until 2005 there was no clear uniform international guideline for this. The number of COPD subjects participating in a study can strongly differ depending whether we use ERS criteria or ATS criteria and this knowledge makes it difficult to compare similar study with each other [21].
Most COPD studies use different criteria of reversibility in COPD like the ERS criteria of Siafakas et al as mentioned in 1995 with only change < 10 % of predicted value after short acting bronchodilator, or the ATS criteria from 1994 and 2005 which prefer a change of > 12% of initial value and > 200 ml after short acting bronchodilator. With the ATS and ERS criteria from 2005 which also prefer a change of > 12% of initial value and > 200 ml after short acting bronchodilator, we can define reversibility in COPD [3] .
Most studies assessing acute effect on FEV_1 in patients with COPD from short-term bronchodilators, showed similar or larger response on anticholinergics compared to that on beta-2-agonists [22-24]. However, other studies have demonstrated equal effect with ipratropium and salbutamol. The reason for a higher anticholinergic response could be increased anticholinergic – mediated muscle tone and reduced mucus secretion [25].
In our study we found that the mean VAS score after bronchodilators and placebo showed significant improvements, but did not significantly differ between each other. Our study could not confirm the results of the study

of Taube et al., in which with regard to the ILP, a strong correlation was found in changes in FIV_1 and dyspnoea VAS score after administration of salbutamol. The lack of correlation in changes of VAS score after acute bronchodilation and the other lung function parameters in our study may be due to the fact that at least forty-four percent of our COPD subjects reached the FIVC within a second. This may cause the FIV_1 to be less sensitive compared to the study of Taube et al[12].

Other differences between the study of Taube and our study might be: firstly, that our spirometric standard differs from his study. Unlike in our study, he used at least two measurements of forced inspiratory and expiratory volumes. Secondly, the population of his study, differs from that of our project. Our patients had probably a more stable COPD due to

the inclusion criteria of the study. We included only subjects who did not have any exacerbation of their COPD within the last two months, did not use systematically corticosteroid or antibiotics in this period and none of our patient used (long-term) oxygen or theophylline. Thirdly, the population of both studies could have different severity of COPD subjects. In our population in 63 patients with COPD the FEV_1 mean was 1.49 ± 0.73 litres, respectively 51 ± 20 % predicted whereas in the study of Taube in 61 patients with COPD the FEV_1 mean was 1.09 ± 0.49 litres, respectively FEV_1: 38 %; range: 13 to 80% predicted. This suggests that in his study the subjects had a more severe COPD. Furthermore his study is limited to the response on salbutamol and lacks the effect on anticholinergic drugs. Additionally, the perception of dyspnoea in subjects can be divided in less or more change in shortness of breath after bronchodilator, low-perceivers respectively high-perceivers [26]. It is possible that in our study we had more low-perceivers. The question arises whether in subjects with stable COPD at rest and a lung function according to GOLD class I and II, relevant changes in dyspnoea after bronchodilators in a short period of 30 minutes may be expected. Noseda et al found a modest relationship ($r = 0.51$) between the change in VAS and FEV_1 in patients with COPD. Other studies report significant correlations with FEV_1 and the baseline and transitional dyspnoea indices (BDI/TDI) of Mahler in patients with COPD [27,28]. We used the same VAS score as Taube but we found no results like in his study. The question arises if the relationship of changes in FEV_1 and dyspnoea score BDI/TDI is better than with VAS score and what the relationship between ILP and BDI/TDI will be.

In conclusion, short-acting bronchodilators in subjects with stable COPD showed significant improvements in both FEV_1 and ILP. ILP were not more sensitive than FEV_1 for detecting responders on bronchodilators. We were not able to reproduce the high degree of correlations Taube et al. found between inspiratory parameters and VAS dyspnoea score.

Reference List

1. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2009. Available from: http://www.goldcopd.org.

2. Celli BR, MacNee W: Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004, 23:932-946.

3. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J et al.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26:948-968.

4. O'Donnell DE, Lam M, Webb KA: Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999, 160:542-549.

5. Belman MJ, Botnick WC, Shin JW: Inhaled bronchodilators reduce dynamic hyperinflation during exercise in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1996, 153:967-975.

6. Anthonisen NR, Wright EC: Bronchodilator response in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1986, 133:814-819.

7. Tantucci C, Duguet A, Similowski T, Zelter M, Derenne JP, Milic-Emili J: Effect of salbutamol on dynamic hyperinflation in chronic obstructive pulmonary disease patients. *Eur Respir J* 1998, 12:799-804.

8. Ramirez-Venegas A, Ward J, Lentine T, Mahler DA: Salmeterol reduces dyspnea and improves lung function in patients with COPD. *Chest* 1997, 112:336-340.

9. O'Donnell DE, Lam M, Webb KA: Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998, 158:1557-1565.

10. O'Donnell DE, Revill SM, Webb KA: Dynamic hyperinflation and exercise intolerance in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001, 164:770-777.

11. O'Donnell DE, Forkert L, Webb KA: Evaluation of bronchodilator responses in patients with "irreversible" emphysema. *Eur Respir J* 2001, 18:914-920.

12. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162:216-220.

13. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Recommendations for the measurement of FIV(1) values in chronic obstructive pulmonary disease. *Respiration* 2008, 76:46-52.

14. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P et al.: Standardisation of spirometry. *Eur Respir J* 2005, 26:319-338.

15. Wanger J, Clausen JL, Coates A, Pedersen OF, Brusasco V, Burgos F, Casaburi R, Crapo R, Enright P, van der Grinten CP et al.: Standardisation of the measurement of lung volumes. *Eur Respir J* 2005, 26:511-522.

16. Quanjer Ph.H., Tammeling G.J.: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993, 6:5-40.

17. Hadcroft J, Calverley PM: Alternative methods for assessing bronchodilator reversibility in chronic obstructive pulmonary disease. *Thorax* 2001, 56:713-720.

18. Visser FJ, Ramlal S, Pelzer B, Dekhuijzen PN, Heijdra YF: Random variation of inspiratory lung function parameters in patients with COPD: a diagnostic accuracy study. *BMC Pulm Med* 2010, 10:28.

19. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1:307-310.

20. Noseda A, Schmerber J, Prigogine T, Yernault JC: Perceived effect on shortness of breath of an acute inhalation of saline or terbutaline: variability and sensitivity of a visual analogue scale in patients with asthma or COPD. *Eur Respir J* 1992, 5:1043-1053.

21. Siafakas NM, Vermeire P, Pride NB, Paoletti P, Gibson J, Howard P, Yernault JC, Decramer M, Higenbottam T, Postma DS et al.: Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. *Eur Respir J* 1995, 8:1398-1420.

22. Braun SR, McKenzie WN, Copeland C, Knight L, Ellersieck M: A comparison of the effect of ipratropium and albuterol in the treatment of chronic obstructive airway disease. *Arch Intern Med* 1989, 149:544-547.

23. Friedman M: A multicenter study of nebulized bronchodilator solutions in chronic obstructive pulmonary disease. *Am J Med* 1996, 100:30S-39S.

24. Passamonte PM, Martinez AJ: Effect of inhaled atropine or metaproterenol in patients with chronic airway obstruction and therapeutic serum theophylline levels. *Chest* 1984, 85:610-615.

25. Barnes PJ: Neural control of human airways in health and disease. *Am Rev Respir Dis* 1986, 134:1289-1314.

26. Noseda A, Schmerber J, Prigogine T, Yernault JC: How do patients with either asthma or COPD perceive acute bronchodilation? *Eur Respir J* 1993, 6:636-644.

27. Casaburi R, Mahler DA, Jones PW, Wanner A, San PG, ZuWallack RL, Menjoge SS, Serby CW, Witek T, Jr.: A long-term evaluation of once-daily inhaled tiotropium in chronic obstructive pulmonary disease. *Eur Respir J* 2002, 19:217-224.

28. Mahler DA, Donohue JF, Barbee RA, Goldman MD, Gross NJ, Wisniewski ME, Yancey SW, Zakes BA, Rickard KA, Anderson WH: Efficacy of salmeterol xinafoate in the treatment of COPD. *Chest* 1999, 115:957-965.

7

The optimization of the diagnostic work-up in patients with suspected obstructive lung disease

Frank J. Visser, Milena J.M.M. van der Vegt, Gert Jan van der Wilt and Julius P. Janssen

BMC Pulm Med. 2010 Nov 23;10:60.

Abstract

Background: Pulmonary function testing is a key procedure in the work-up of patients who are suspected of having asthma and chronic obstructive lung disease (COPD). Therein, clinical visits and pulmonary function tests (PFTs) are the major contributors to the overall financial costs.
The aim of this study was to assess whether a specific diagnostic test protocol contributes to the optimization of the work-up of patients who are suspected of having asthma and COPD.

Methods: A prospective, single-blind, and randomized controlled study was performed. In the control group (CG), all of the PFTs that were ordered by the lung specialist were carried out. In the experimental group (EG), specific PFTs were selected according to our protocol. The primary end point was the total cost of achieving a final diagnosis.

Results: One hundred and seventy-nine patients were included into this study: 86 in the CG and 93 in the EG. The mean number of tests to diagnosis was 3.8 in the CG versus 2.9 in the EG ($P<0.001$). The mean number of redundant PFTs before diagnosis was 1.2 in the CG versus 0.08 in the EG ($P<0.001$). The number of patients who required an additional outpatient visit to complete diagnosis was higher in the CG in comparison to the EG ($P=0.02$). The mean cost of work-up per diagnosis was €227 in the CG versus €181 in the EG ($P<0.001$).

Conclusions: In this group of patients with suspected obstructive lung disease, protocol-driven, PFT-based selection is more cost-effective than test selection at the discretion of lung physicians.

Key words: COPD, asthma, cost-effectiveness, pulmonary function tests, diagnostics, redundant tests

Introduction

Diagnosing asthma and COPD is an important part of the daily practice of pulmonary physicians. Pulmonary function tests (PFTs) play a key role in the work-up of obstructive pulmonary diseases [1-3].

No exact figures exist for the annual costs that are associated with current diagnostic processes, although they are likely to be substantial.
Finding the optimal diagnostic work-up in patients with obstructive lung disease is challenging. A physician who routinely orders most or all PFTs in the work-up of asthma and COPD patients, runs the risk of unnecessary testing; however, a physician who orders tests more sparingly, runs the risk of unnecessary outpatient visits. In view of the high incidence of patients with obstructive lung diseases, it is important to find the optimal diagnostic work-up in each of these patients. To this end, we have developed a diagnostic protocol (Figure 1) that can be jointly used by physicians and pulmonary function assistants. In our group, some physicians already use this diagnostic PFT protocol; however, some of the physicians order PFTs without following the prescribed diagnostic PFT protocol.

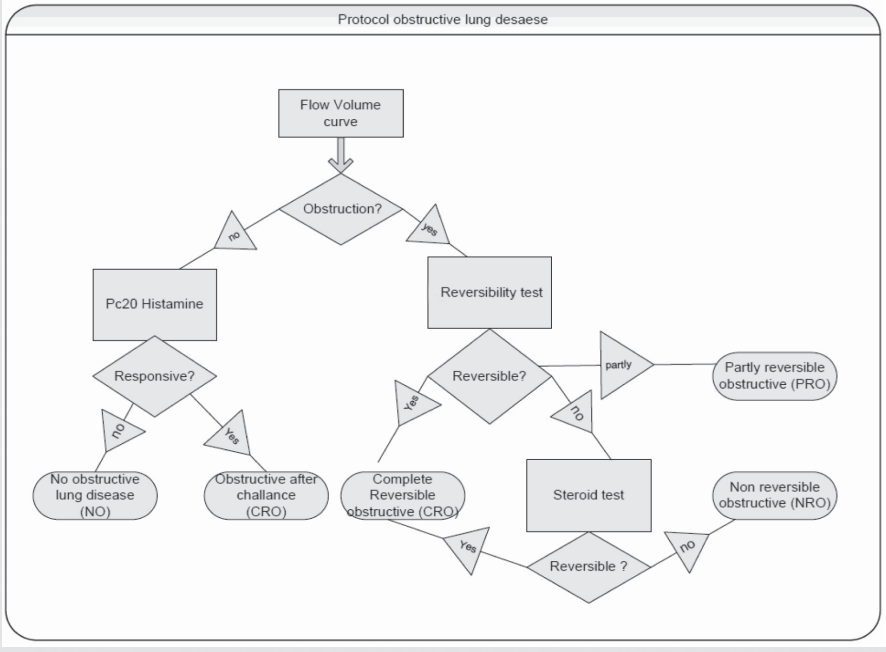


Figure 1. *Pulmonary function protocol for obstructive diseases. Criteria for obstruction, airway responsiveness (PC20 histamine), reversibility, and steroid tests (see text).*

Prior to the beginning of this study, there was no available evidence that demonstrated that protocol-driven PFT ordering is more efficient than physician-driven test ordering. Therein, we hypothesize that protocol-driven test ordering will be more efficient than test ordering without direction from a diagnostic protocol.

The aim of this study was to assess whether protocol-driven test ordering reduces the number of redundant pulmonary function tests, decreases the number of outpatient visits, and increases the cost effectiveness of patient work-up in comparison to physician-driven test ordering.

Methods

A prospective, randomized, and single-blind trial was conducted at our outpatient unit.
An institutional review board (IRB) approved this study. This study was only a formal stratification of the current practice; hence, informed consent was not necessary.
The study participants consisted of consecutive adult patients, who were referred to our respiratory outpatient clinic and suspected to have asthma or COPD at the end of the first outpatient visit. None of the patients had recently (in the preceding three years) been diagnosed with asthma or COPD by a pulmonary physician. Patients were primarily referred by general practitioners; however, two patients referred themselves to our outdoor department, and three were referred by cardiologists. We excluded patients who were not able to adequately complete pulmonary function tests, were referred to a pulmonary physician because of an abnormal X-ray, needed pre-operative consultations, and have had an infection of the upper or lower airways or a possible exacerbation of obstructive airway disease in the past two months.
In the first visit, the physician takes a medical history, performs a physical examination, makes a differential diagnosis, and orders laboratory testing, such as a chest X-ray. Only patients who were most likely to be diagnosed with asthma or COPD were included into this study. Physicians ordered diagnostic tests as they deemed appropriate and added the reason for the pulmonary function testing (e.g. suspected obstructive lung disease). At the end of the first outpatient visit, nurses randomized the patients into the control

group (work-up at the discretion of the physician) or the experimental group (work-up in accordance with the protocol) by pulling an opaque envelope. PFT lab assistants were notified of the outcome of the randomization in order to allow them to perform investigations as ordered or per protocol.

The PFT protocol is shown in Figure 1. Therein, when the forced vital capacity (FVC) is less than normal, a total lung capacity (TLC) measurement was conducted in order to exclude restrictive lung disease. When the patients smoked more than 10 pack years, their respective diffusion capacities were measured.

A second physician independently examined the results for each patient and classified the patients as follows:

Completely reversible (CR): (1) An airway obstruction that is completely reversible (FEV1 reversible by ≥ 9% of the predicted value) to a normal range after beta-2 agonist and anticholinergic treatment; patients who received reversibility testing were tested for both bronchodilators; (2) An airway obstruction that is completely reversible after 14 days of 30 mg/day of prednisone; and (3) normal PFTs but a decreased PC20 histamine threshold.

This PFT group supports to the diagnosis of asthma.

Non-reversible obstructive (NRO): Reduced FEV1 and FEV1/FVC values, which are irreversible after beta-2 agonist and/or anti-cholinergic treatment (an FEV1 increase of < 9% of the predicted value) and no return to the normal range after 10 days of 30 mg/day of prednisone.

This PFT group supports the diagnosis of COPD.

Partly Reversible Obstructive (PRO): Reversibility is present, but an airflow limitation persists. FEV1 increases by ≥ 9% of the predicted normal value but does not return to a normal range after bronchodilators.

This PFT group supports the diagnosis of asthma with persistent airflow limitation or COPD with partly reversibility.

Normal PFT group: No airflow limitation and a normal PC20 histamine threshold.

This group does not support the diagnosis of obstructive lung disease; hence, a different diagnosis must be considered.

The second physician assessed if the appropriate tests were conducted according to the diagnostic flow and decided on the pulmonary function classification. He assessed the decision about the final diagnosis of the PFT-referring physician.

The second physician calibrated his findings with the findings of the PFT ordering physician only if there were conflicting findings. In all of these cases, we achieved a consensus on the final diagnosis.

COPD and asthma were finally diagnosed by the first physician on the basis of medical history (smoking behavior, allergies, a family history of asthma, and/or a pre-existing childhood condition); PFTs and clinical investigations, such as the eosinophil count; and the radioallergosorbent test (RAST).

In the protocol, the criteria for obstruction included an FEV1 < normal and an FEV1/FVC < normal according to Quanjer et al. [4]. Airway hyperresponsiveness was defined as PC20 histamine < 4 mg/ml [5]. Reversibility was defined as a ≥9% improvement in the FEV1 in comparison to a predicted normal value [6-8]. Steroid tests consisted of 30 mg/day of prednisone for 10 days with the intent of reversing the FEV1 to normal levels, as advised by the Dutch Committee: diagnosis for asthma and COPD [7].

Forced vital capacity (FVC), forced expiratory volume in one second (FEV1), and airway responsiveness (PC20 Histamine) were measured according to ERS criteria [1].

During the follow-up visit, the results of all of the investigations that were carried out at the discretion of the physician or according to the protocol were available to the physician who then decided whether a final diagnosis could be made. Follow-up visits and additional PFTs were scheduled as deemed appropriate.

Redundant PFTs were defined as tests that were not absolutely necessary to establish a final diagnosis. For example, a reversible obstructive PFT made the histamine provocation test redundant, whereas a normal flow volume curve made the reversibility test redundant.

The economic analysis was conducted from a health care perspective that included only direct medical costs. Where available, unit cost prices were derived from a national guideline for the economic analysis of health care services [9]. In other instances, real cost prices were calculated on the basis of hospital administration data (Table 1).

.

Table 1: *The cost of PFTs and follow-up visits*

Tests	Cost
Flow volume curve	€15,00
Reversibility (bronchus dilators) testing	€18,00
TLCO (diffusion capacity)	€41,00
Hyper reactivity (histamine)	€84,00
TLC	€41,00
Outpatient visit	€41,00
Steroid test	€33,00

Statistical analysis

The unpaired t-test with Welch's correction was used to test for the statistical significance of differences between the two groups. An alpha of 0.05 or less was considered to be significant. For statistical calculations, we used GraphPad Prism5 for Microsoft Windows (www.graphpad.com).

Results

From a total of 183 patients, 179 patients were included in this study: 86 patients in the CG and 93 patients in the EG. Four patients were excluded for the following reasons: one patient had a malignancy, one patient failed to follow-up, one patient died within a week of the start of the study, and the protocol was not followed with one patient. The second physician calibrated his findings with the findings of the PFT-ordering physician only if there were conflicting findings. In all of these cases, we achieved a consensus on the final diagnosis.

Table 2 summarizes patient characteristics at baseline. Classification of PFT groups on the basis of PFT results and final diagnoses are presented in Table 3. In the control group, 35 patients were sorted in the completely reversible obstructive (CRO) group versus 39 patients in the experimental group. For the non-reversible obstructive group, these numbers were eight and nine patients, respectively. For the partly-reversible obstructive group, these numbers were 12 and 14, respectively.

***Table 2**. Patient characteristics*

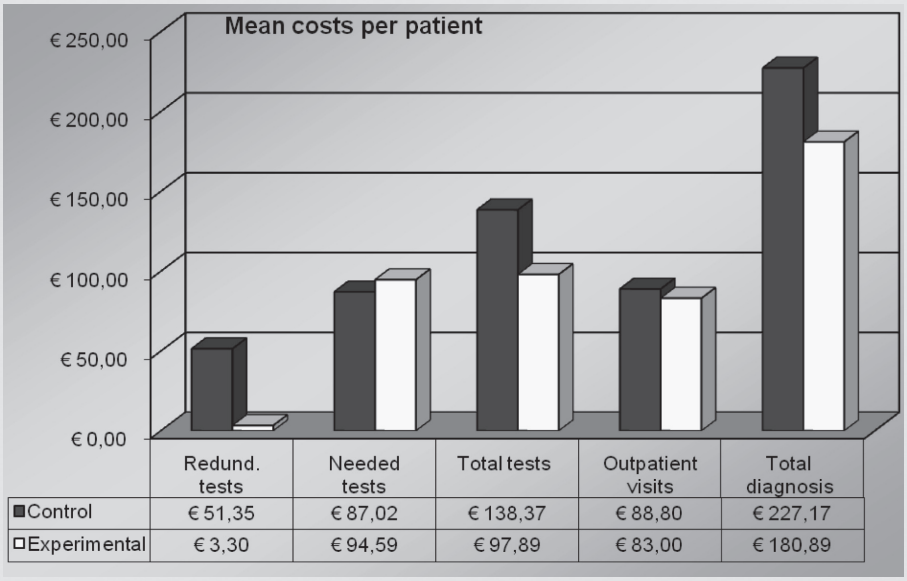
	Control group	Experimental group
Number of patients	86	93
Avg. age (years)	55.4	56.8
Sex (% male)	45	46
Height	168.4	170.8
FEV1 % pred.	77.8	78.0
FVC % pred.	91.3	91.0
Non-smoker	41%	38%
Ex-smoker	20%	23%
Current smoker	33%	29%
Unknown smoker	6%	10%

***Table 3.** Classification on the basis of PFT results and final diagnoses*

Classification on the basis of PFT	Control group	Experimental group
Completely reversible obstructive (CRO)	41%	42%
Non-reversible obstructive (NRO)	17%	17%
Partly reversible obstructive (PRO).	14%	15%
Normal PFT (NO)	27%	26%
Final diagnosis		
Asthma	41%	42%
Asthma partly-reversible	5%	1%
COPD non-reversible-	17%	17%
COPD partly-reversible	10%	14%
Other diagnosis	27%	26%

Diagnosis in the normal PFT group consisted of sarcoidosis (1), gastroesophageal reflux (GER) (5), rhinitis and/or sinusitis (7), hyperventilation syndrome (3), persistent cough (15/21 were smokers or ex-smokers), and dyspnea (3/8 were smokers).
On the basis of clinical assessment, 52 patients were found to have COPD with the following classifications: 7 patients in GOLD stage 1; 24 patients in stage 2; 13 patients in stage 3, and 8 patients in stage 4.

The total cost of the procedures that were used to reach a final diagnosis in patients who were suspected of having obstructive lung diseases.
In the control group, the mean total cost of testing and outpatient visits per diagnosis were €227.17, whereas, in the experimental group, this cost was €180.89 (Figure 2, P<0.001), which is a 20% reduction in cost.



	Redund. tests	Needed tests	Total tests	Outpatient visits	Total diagnosis
Control	€ 51,35	€ 87,02	€ 138,37	€ 88,80	€ 227,17
Experimental	€ 3,30	€ 94,59	€ 97,89	€ 83,00	€ 180,89

Figure 2. *Mean cost per patient per diagnosis. Redund.= Redundant.*

The number and cost of outpatient visits until diagnosis
In the control group, two outpatient visits were needed to reach a final diagnosis in 71 patients, whereas 15 patients needed three outpatient visits (mean 2.17, median 2). In the experimental group, 90 patients had two visits, and 3 patients had three visits (mean 2.03, median 2). The difference in total visits between these groups was statistically significant (P=0.02). Mean costs were €88.80 in the control group and €83.00 in the experimental group (Figure 2, P=0.02).

The number and cost of the PFTs that were needed for diagnosis:
In the control group, a mean number of 3.81 PFTs per patient was necessary in order to diagnose asthma or COPD. In the experimental group, a mean number of 2.94 tests was necessary (Figure 2, P<0.001). In the control and experimental groups, the mean costs of PFTs were €138.37 and €97.89, respectively (P<0.001).

The number and cost of the redundant PFTs that were used for diagnosis:
In the control group, a mean of 1.20 redundant PFTs per patient were performed. In the experimental group, a mean of 0.08 unnecessary PFTs were performed (P<0.001). The mean costs of redundant tests for diagnoses in the control and experimental groups were €51.35 and €3.30, respectively (Figure 2).
Histamine provocation tests and the added cost of reversibility testing were the two most important sources of redundant costs (Table 4).

Table 4. *Redundant pulmonary function tests (PFT).*

Redundant PFT	Control group		Experimental group	
	Number (%)	Cost in €	Number (%)	Cost in €
Histamine provocation test	19 (22)	1596	1 (1)	84
Reversibility testing	30 (35)	540	1 (1)	18
TLC	30 (35)	1230	3 (3)	123
Diffusion capacity for CO	24 (28)	984	2 (2)	82
Totals	103 (120)	4323	7 (8)	307

Time until final diagnosis
In the control group, a mean number of 33.02 days was necessary to reach a final diagnosis, whereas in the experimental group, a mean number of 35.94 days was needed (P=0.51).

Post-hoc
We evaluated the added value in making a diagnosis of asthma using the steroid test in our patients. Conforming to the protocol, we needed 11 steroid tests in the control group and 10 steroid tests in the experimental group. Patient diagnoses did not change with the addition of these steroid tests.

Discussion

The main finding of our study is that the introduction of a problem-oriented protocol for ordering PFTs in patients with suspected obstructive pulmonary disease can reduce the number of redundant PFTs and outpatient visits, which results in a 20% decrease in costs without an increase in time to final diagnosis. Given the high frequency of PFT usage for the diagnosis of obstructive lung disease, this observed decrease in cost results in a substantial savings at a population level. In our practice of 600 patients per year with suspected obstructive lung disease, protocol-guided test ordering can lead to an annual cost reduction of €27,768 = €64.28 per patient. The most important part of these potential savings is a reduction in the need for reversibility testing when a normal flow volume curve is obtained and a reduction in the need for the time-consuming and unpleasant hyperresponsiveness (PC20 histamine) test when obstruction with reversibility is obtained.

Our lung function protocol is based on the asthma and COPD guidelines [7] of the Netherlands and is within the ERS and ATS standards. [3,4,6,10] Therefore, many other countries can use our protocol with slight modifications. Our patients with asthma and COPD are demographically similar to patients in other western European countries, the USA, and Canada. The only difference is that most patients were referred to us from a family doctor, as is typically the case in the UK; however, in some other countries, a family physician may be skipped more often. Therefore, a minor selection bias is possible; however, the diagnostic criteria for COPD or asthma do not depend on the patient's physician. Of course, the skills and the tools are different between general practitioners (GPs) and pulmonary physicians; however, for this study, we included only the most basal lung function tests and omitted tests, such as exercise testing. These tests are not always needed to confirm a diagnosis of asthma or COPD. We believe that the protocol discussed herein could easily be followed by GPs or hospital physician assistants so long as they have access to these basic tests, thus, leading to potentially more health-care savings.

As we stated before, some physicians have used the protocol-driven lung function protocol that we developed approximately three years prior to this study. The first author of this study and two other physicians have primarily followed the protocol-driven testing strategy. Two other physicians (one

senior and one junior) did not implement this protocol-driven testing strategy because they were not convinced of the ability of this protocol to save time and cost.
We all agreed to perform this study, and the behavior of these two physicians did change after the completion of this study.
Trainees stay for four years in our hospital, and, at the time of the study, we had five trainees in all stages of their educational processes. We asked our trainees to order according to a test-protocol rather than at their own discretion. We advised them that this would be more efficient.
The protocol allows pulmonary function assistants to work more efficiently, which decreases the frequency that they need to interrupt the doctor, who will often be in a consultation. The workflow is based on our national guideline, which resembles the international guidelines of the ERS.

Steroid tests would not have added value to the diagnostic workflow in our patient group; hence, we doubted the need for such a test in a routine setting. After this study, we removed the steroid tests from the PFT protocol.
The total time in days to diagnosis was not different between the two groups; however, without waiting lists for the outpatient department and for PFT, we believe there will be a difference in time between the CG and the EG that favors the EG.
We want to emphasize that international guidelines and several national guidelines do not recommend reversibility testing as a means to distinguish asthma from COPD, other than when lung function returns to normal limits. We used three pulmonary function groups as an intermediate; however, the final diagnosis of asthma or COPD (or both) can only be made when the full clinical context, in which PFTs are only a part, is considered.
The prediction threshold of 9% is not a commonly accepted threshold for distinguishing asthma from COPD. We only use this criterion to distinguish between the “non-reversible obstructive” and “partly-reversible obstructive” PFT groups. Internationally, this is merely one method that can be used to make this distinction between these subgroups, and no consensus is available regarding which criterion is the best [6].
A weakness in our study is the potential of a Hawthorne effect. The knowledge that they were involved in a trial may have affected the physicians’ behavior (test ordering, making a final diagnosis). The only way to avoid this problem is to retrospectively conduct the study, which would challenge its internal

validity. Therefore, we decided to conduct a random, parallel design so as to ensure internal validity (equivalence between groups, minimal likelihood of confounding); however, to the extent that our study suffered from a potential Hawthorne effect, this will most likely have resulted in an under-estimation of the impact of protocol-guided test ordering.
In order to minimize the impact of a potential Hawthorne effect, we discussed all of the results within our group after the inclusion period. Some doctors wanted to minimize patient return visits and ordered tests, which were regarded to be unnecessary during the second visit.

Conclusion

Problem-orientated PFT ordering, significantly reduces the number of PFTs, the total cost, and the number of outpatient visits in the diagnosis of asthma and COPD.

Reference List

1. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A *et al*.: Standardisation of spirometry *Eur Respir J* 2005, 26: 319-338.

2. Siafakas NM, Vermeire P, Pride NB, Paoletti P, Gibson J, Howard P *et al*.: Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. *Eur Respir J* 1995, 8: 1398-1420.

3. Global Initiative for Asthma (GINA). Global strategy for asthma management and prevention (updated 2005): . 2006. Ref Type: Internet Communication

4. Quanjer Ph.H., Tammeling G.J.: Standardized lung function testing. Official statement of the ERS. *Eur Respir J* 1993, 6: 5-40.

5. Cockcroft DW: Bronchoprovocation methods: direct challenges. *Clin Rev Allergy Immunol* 1994, 24: 19-26.

6. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R *et al*.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26: 948-968.

7. Commissie diagnostiek astma bronchiale en COPD N: Standaard diagnostiek van astma bronchiale en COPD. *Pulmoscript* 1994, *1994;* 31-46.

8. DalesRE STS, SuissaS: Clinicalinterpretationofairwayresponsetoabronchodilator :epidemiologicconsiderations. *Am Rev Respir Dis 1988*, 138: 317-320.

9. Oostenbrink JB, Bouwmans CAM, Koopmanschap M.A., Rutten FFH: Maual for costing studies. Methods and standard cost prices for economic evaluations in health care. *National Health Insurance Board (in Dutch)* 2004.

10. Lung function testing: selection of reference values and interpretative strategies:oficial statement ATS. *Am Rev Respir Dis 1991*, 144: 1202-1216.

8

Summary

Summary

Basal lung function tests are obligatory to diagnose asthma and COPD as can be derived by their definitions:
" (COPD) is a preventable and treatable disease characterized by airflow limitation that is not fully reversible. The airflow limitation is usually progressive and is associated with an abnormal inflammatory response of the lungs to noxious particles or gases, primarily caused by cigarette smoking. Although COPD affects the lungs, it also produces significant systemic consequences [1]."
"Asthma is a disorder defined by its clinical, physiological and pathological characteristics.
The main physiological feature of asthma is episodic airway obstruction characterized by expiratory airflow limitation.[2]"
Forced expiratory volume in the first second (FEV1) may be the most important lung function parameter and is used as a measure of severity of airflow obstruction. It is however disappointing that this parameter does not correlate well with the dyspnea feeling of the patient. Arguments for this finding are:
The weak association between the change in dyspnea (one of the major complaints in subjects with COPD) and the change in FEV1 or FVC after the use of bronchodilators and may be due to airway compression during forced expiration [3-7]
The absence of dyspnea at rest in patient with stable COPD, and therefore an inability to improve after the use of bronchodilators.
Inspiratory parameters, however, are not influenced by the above mentioned airway compression. Therefore we are interested in inspiratory parameters, in subjects with COPD. Inspiratory parameters could be more sensitive to interventions like bronchodilators because the airway collapse may obscure benefits of interventions on expiratory parameters. In the literature O'Donnell and Taube [7,8] give support for the idea that inspiratory parameters may be more sensitive to find significant improvements after bronchodilators in patients with COPD. However, which method of obtaining inspiratory parameters would be best is not defined in the literature.
Because inspiratory parameters can be obtained after a fast expiration (Fe) or after a slow expiration (Se) we compared these methods in chapter two [9].In 169 patients with COPD, Forced Inspiratory volume in the first

second (FIV1) maneuvers were performed after a forced (FIV1-Fe) and a slow (FIV1-Se) expiration. We found that the variability of FIV1-Fe was greater than that of FIV1-Se. The mean difference between FIV1-Se and FIV1-Fe was 0.21 liters. The difference depended on the Global Initiative for Chronic Obstructive Lung Disease (GOLD[1] stage): the higher the GOLD stage, the greater the difference between the two techniques. The correlation coefficient between FIV1-Se and FIV1-Fe was high, but there was a poor agreement between measurements within the same subject of these parameters (limits of agreement -0.52 to +0.94 liters). We concluded that the two methods measuring FIV1 may not be interchangeable and we prefer the FIV1-Se method because the variability was lower.

After selecting the method (FIV1-Se), we investigated the optimal number of the FIV1-Se manoeuvres out of 8 attempts and found that at least 5 FIV1-Se manoeuvres are needed to get an acceptable FIV1. This number holds for all GOLD stages[9].

Apart from the importance of the FEV1 to measure severity of COPD, it is also the most frequently used parameter to measure hyperresponsiveness and reversibility after an intervention. Bronchodilator drugs are widely used in COPD patients and the effect of bronchodilators can be assessed by measuring the acute response (reversibility) of the drug of interest on FEV1. In samples of patients with COPD there are significant improvements in FEV1 both on beta mimetic as on parasympathicolytic drugs, but in an individual patient it is hard to prove improvement based on the recommendations of the ERS-ATS taskforce for interpreting the response of bronchodilators on FEV1 [10]. Only about 11% [11,12]of patients with severe COPD responded when we followed the ERS-ATS criteria, being an improvement of at least 200 ml and 12% from initial value. In chapter three we explored these ERS-ATS criteria. We evaluated the literature and found that none of the referenced literature for these criteria, used the recommended method by Bland and Altman [4] for describing natural spread. Notably this Bland and Altman method was also recommended by the ATS-ERS taskforce [10]. The knowledge of the natural spread is important because only responses to bronchodilators beyond the natural spread or random variation (RV) can be attributed to the bronchodilator in an individual patient. We investigated whether the type of this natural spread (or RV) for FEV1 was homoscedastic or heteroscedastic. A homoscedastic scatter would reduce the chance of under- or overestimating the RV for low or high parameter values. Not

only the type of scatter but also the quantity of scatter expressed in the coefficient of repeatability (CR) is important. In our sample of 79 stable COPD patients, the FEV1 was measured five times in one day. FEV1 values taken within one hour were compared, the difference between the FEV1 measurements was expressed as absolute change in liters or a relative change in percent of predicted value or percent of initial value. We found that the type of scatter was heteroscedastic for the FEV1 when the difference was expressed as an absolute value in liters. However, when the difference was expressed as the percent change from the initial value or predicted value, we found a homoscedastic scatter. The CR within one hour of the FEV1 was 12% when expressed as the percent change from the initial value. Based on these findings, we conclude that the absolute 200 ml=0.2 liter criterion may be omitted and so more patients with severe COPD will have significant airway responsiveness.

In Chapter 2, we already investigated which inspiratory maneuver would be best, now we can use this method to investigate the random variation (natural variability) of the inspiratory parameters. In chapter 4 we investigated the random variation (type and quantity) of inspiratory parameters in the same population as we did for the FEV1[13].
The forced inspiratory volume in 1 second (FIV1), inspiratory capacity (IC), maximal inspiratory flow at 50% (MIF50) and peak inspiratory flow (PIF) were measured five times in one day. The values of these parameters, taken within one hour and within one day were compared. The coefficient of repeatability (CR) was calculated and in addition, linear regression was performed to investigate the type of scatter (homo- or heteroscedastic) of the measured parameters. The type of scatter was heteroscedastic for all of the parameters when the differences were expressed as absolute values; however, when the differences were expressed as the percent change from the initial values, we found a more homoscedastic scatter like we found for the FEV1. The CR within one hour of each parameter expressed as the percent change from the initial value was: IC, 19%; FIV1, 14%; PIF, 18%; MEF50, 21%.
Now we have estimated the random variation for expiratory and inspiratory parameters, it is opportune to measure changes in these parameters on several interventions, like short acting bronchodilators, histamine, pursed lip breathing and steroids. These interventions are also subject for another

thesis, except the pursed lips breathing which we described in chapter 5 [14] and the short acting bronchodilators discussed in chapter 6 [15].

In patients with severe chronic obstructive pulmonary disease (COPD), pursed-lips breathing (PLB) improves the pulmonary gas exchange and hyperinflation measured by electro-optic coupling. The response to PLB in inspiratory lung function tests is not known yet. In chapter 5 we measured the effect of PLB on inspiratory parameters in thirty-five subjects with stable COPD and a forced expiratory volume in first second (FEV1) <50% of the predicted value. Patients were tested for the following primary parameters before, immediately after PLB, and 5 min after PLB: forced inspiratory vital capacity, inspiratory capacity (IC), forced inspiratory volume in first second (FIV1), maximal inspiratory flow at 50% of VC, (MIF50) and peak inspiratory flow (PIF). Patients were also tested for secondary parameters: vital capacity (VC), FEV1, breathing frequency (BF), end-tidal CO2 tension (ET-CO2), and oxygen saturation (SO2).
Of the primary parameters, only IC improved significantly (89 ml p=0.006); with regard to the secondary parameters, the mean SO2 improved by 1% (p = 0.005) and the mean ET-CO2 and BF decreased significantly to 3.2 mm Hg and 3.1 breaths/min, respectively. After 5 min the effects diminished.

In chapter 6 the responsiveness of short-term bronchodilators were described on inspiratory lung function parameters (ILP),and FEV1[15].
The aim of this study was to assess the effects of inhaled salbutamol 400 mcg, ipratropium 80 mcg and placebo on ILP and FEV1, and their relation with change in dyspnea feeling on a Visual Analogue Scale (VAS).
85 subjects with stable COPD participated in a cross-over, randomized, double-blind placebo-controlled study. Spirometry was performed before and after inhalation of salbutamol, ipratropium and placebo. Changes in dyspnea were measured with VAS.
All ILP and FEV1 improved significantly on bronchodilators, except FIF50 after Ipratropium. After both bronchodilators, percent changes from initial values in IC were higher than corresponding changes in FEV1, but the differences were not significant (all p>0.06).
Mean VAS score after bronchodilators and placebo showed significant improvements in dyspnea feeling but did not significantly correlate with changes in lung function parameters. For each lung function parameter,

patients were further classified as responder (improvement beyond random variation) or non-responder according to the degree of the change from baseline. Response rates did not significantly differ between the various ILP. Also no significant differences were found between responders and non- responders regarding dyspnoea after bronchodilators. This applied to all ILP as well as to FEV1. So in subjects with COPD, all ILP and FEV1 showed, just like the VAS score, significant improvements after bronchodilators, adjusted for placebo effect. However ILP were not more sensitive than FEV1 in detecting responders to bronchodilators and for changes in dyspnea feeling.

Asthma and COPD are very common and lung function tests are mandatory to establish their diagnosis. Therefore, it is important that lung function tests are effective and efficient used,since pulmonary function tests (PFTs), together with clinical visits, are the major contributors to the overall financial costs of diagnosing patients suspected of having obstructive lung diseases. In chapter 7 we assessed whether a specific diagnostic test protocol contributes to the optimization of the work-up of patients suspected of having asthma or COPD. A prospective, single-blind, and randomized controlled study was performed. In the control group (CG), all of the PFTs that were ordered by the lung specialist were carried out. In the experimental group (EG), specific PFTs were selected according to a protocol. The primary end point was the total costs of achieving a final diagnosis. One hundred and seventy-nine patients were included into this study: 86 in the CG and 93 in the EG. The mean number of tests to diagnosis was 3.8 in the CG versus 2.9 in the EG ($P<0.001$). The mean number of redundant PFTs before diagnosis was 1.2 in the CG versus 0.08 in the EG. The number of patients who required an additional outpatient visit to complete diagnosis was higher in the CG in comparison to the EG. The mean cost of work-up per diagnosis (including clinical visits and lungfunction tests) was €227 in the CG versus €181 in the EG. In patients with suspected obstructive lung disease, protocol-driven, PFT-based selection is more cost-effective than test selection at the discretion of lung physicians

Discussion

The most important results described in this thesis can be summarized as follows:
Inspiratory parameters could better be assessed after a slow expiration than after a forced expiration.
Five inspiratory maneuvers were needed to get an optimal result.
We found an unwanted heteroscedastic spread for the FEV1 and inspiratory lung function parameters, if we expressed the change in liters or L/sec, however if we expressed the change in percentage form baseline value we found a more preferred homoscedastic spread.
For the FEV1 parameter we did not find evidence for the 200 ml improvement criterion as stated by Pellegrino at al. Just 12% improvement from baseline value is needed for a statistically significant improvement.
Inspiratory parameters were not better than FEV1 in the ability to detect a response beyond the natural spread after short acting bronchodilators.
Pursed lips breathing resulted in an improved inspiratory capacity (a measure for hyperinflation) but not in improved lung function parameters of airflow.
In patients with suspected obstructive lung disease, protocol-driven, PFT-based selection was more cost-effective than test selection at the discretion of lung physicians.

In chapter 6, we concluded that for routine lung function testing there is no need to test the inspiratory parameters extensively. Although inspiratory parameters improved significantly after bronchodilators, FEV1 disclosed more responders (response beyond the random variation), than inspiratory parameters. In our study, the correlation between inspiratory parameters and dyspnea scale was not statistically significant. In the literature, significant correlation between these parameters have been published. This discrepancy may be due to the fact that we selected COPD subjects in all stages of severity whereas Taube et al selected only more severe COPD patients. But even if we selected a subset of severe COPD patients (GOLD stage III and IV), we were not able to detect a significant relation. Another explanation may be that a correlation on dyspnea may be hard to find in subjects with stable COPD because they are mostly not short of breath at rest.
We also found that a lot of subjects with COPD were able to inspire within one second their whole vital capacity, so FIV1 may not be so sensitive at all

in detecting bronchodilator response and another parameter like FIV in half a second would be more appropriate.

Recommendations for further research

Further research is ongoing, including research on the response of inspiratory parameters after inhalational steroids long acting bronchodilators and histamine. During our study more questions arised to make our function protocol more efficient:

can we integrate the fraction of NO in the expired air (FeNO) in the lungfunction protocol, and so diminish the need of histamine provocation tests which are not patient friendly.

for responsiveness testing, is it valid to use salbutamol first and measured the response after 15 minutes immediately followed by ipratropium and measure the added value response 30 minutes later as we did, compared with salbutamol and ipratropium on different days?

For responsiveness testing, is it valid to use salbutamol and ipratropium in one gift and measure the responses of the two after 15 and 45 minutes respectively?

On inspiratory parameters we did not investigate the random variation of the inspiratory capacity before and after maximal voluntary ventilation. This could be important to investigate dynamic hyperinflation that is correlated with dyspnea on exertion.

If we know the above mentioned random variation we are able to study interventions like short acting bronchodilators on the effect on dynamic hyperinflation without the need for exercise testing.

Reference List

1. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2009. Available from: http://www.goldcopd.org.

2. From the Global Strategy for Asthma Mangement and Pprevention, Global initiative for Asthma (GINA) 2010. Available from http://www.ginasthma.com.

3. Archibald CJ, Guidotti TL: Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab* 1987, 1:45.

4. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1:307-310.

5. Cazzola M: Should the choice of a long-acting bronchodilator in the long-term therapy of chronic obstructive pulmonary disease depend entirely on the onset of action? *Respiration* 2006, 73:410-411.

6. Guyatt GH, Townsend M, Nogradi S, Pugsley SO, Keller JL, Newhouse MT: Acute response to bronchodilator. An imperfect guide for bronchodilator therapy in chronic airflow limitation. *Arch Intern Med* 1988, 148:1949-1952.

7. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162:216-220.

8. O'Donnell DE, Forkert L, Webb KA: Evaluation of bronchodilator responses in patients with "irreversible" emphysema. *Eur Respir J* 2001, 18:914-920.

9. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Recommendations for the measurement of FIV(1) values in chronic obstructive pulmonary disease. *Respiration* 2008, 76:46-52.

10. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J et al.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26:948-968.

11. Han MK, Wise R, Mumford J, Sciurba F, Criner GJ, Curtis JL, Murray S, Sternberg A, Weinman G, Kazerooni E et al.: Prevalence and clinical correlates of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 35:1048-1056.

12. Hansen JE: Assessment of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 36:972.

13. Visser FJ, Ramlal S, Pelzer B, Dekhuijzen PN, Heijdra YF: Random variation of inspiratory lung function parameters in patients with COPD: a diagnostic accuracy study. *BMC Pulm Med* 2010, 10:28.

14. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Pursed-lips breathing improves inspiratory capacity in chronic obstructive pulmonary disease. *Respiration* 2011, 81:372-378.

15. Ramlal SK, Visser FJ, Hop WC, Dekhuijzen PN, Heijdra YF: Reversibility of inspiratory lung function parameters after short-term bronchodilators in COPD. *Respir Physiol Neurobiol* 2010, 173:58-63.

9

Samenvatting

Samenvatting

Basaal longfunctieonderzoek en met name spirometrie is noodzakelijk om tot de diagnose astma of COPD (Chronic Obstructive Pulmonary Disease) te komen, hetgeen ook is af te leiden uit de definities van deze aandoeningen: "COPD is een aandoening die men kan voorkomen en behandelen en wordt gekenmerkt door een beperking van de luchtstroom die niet volledig hersteld kan worden. De beperking van de luchtstroom neemt gewoonlijk steeds meer toe en gaat gepaard met een abnormale ontstekingsreactie van de long ten gevolge van schadelijke deeltjes of gassen in de inademingslucht, dit laatste weer vooral als gevolg van het roken van sigaretten. Hoewel COPD primair een aandoening van de longen is, heeft het ook belangrijke consequenties voor andere organen en de psyche [1]."
"Astma is een aandoening die wordt gedefinieerd door klinische, fysiologische, en pathologische kenmerken. Het belangrijkste fysiologische kenmerk is de in episoden (aanvallen) optredende luchtwegobstructie en daardoor beperking van de luchtstroom [2] ".

De geforceerde uitademing in de eerste seconde na een volledige inademing (FEV1) is tot nu toe de belangrijkste longfunctieparameter die wordt gebruikt o.a. als maat voor de ernst van de luchtweg obstructie. Echter stemt de FEV1 weinig overeen met het gevoel van benauwdheid dat de patiënt met astma of COPD ervaart. Argumenten hiervoor zijn:

1. Bij patiënten met COPD is er een zwakke overeenstemming gevonden, voor en na het toedienen van luchtwegverwijders, tussen de verbetering in het gevoel van benauwdheid en de gemeten verbetering van de longfunctie en met name de FEV1 [3-7]. Dit is mogelijk te wijten is aan de luchtwegcompressie gedurende de geforceerde uitademing (expiratie).

2. Bij patiënten met stabiel COPD, die per definitie een te lage FEV1 hebben, komt de ervaring van kortademigheid in rust weinig voor. Kortademigheid kan dan ook niet verbeteren na inname van een luchtweg (bronchus) verwijdend medicament of een medicament dat de, bij ontsteking horende, zwelling (oedeem) vermindert, terwijl de FEV1 dan wel verbetert. Veel studies waaronder de onze werden verricht bij stabiele COPD patiënten.

Spirometrische parameters bij de inademing (inspiratie) worden niet beïnvloed door de onder punt 1 hierboven genoemde luchtwegcompressie, maar kunnen wel beïnvloed worden door vernauwing ten gevolge van ontstekingsoedeem of samentrekken van de spieren in de wand van de luchtwegen. Juist daarom zijn wij geïnteresseerd in de zogenaamde inspiratoire parameters met name bij patiënten met COPD. Inspiratoire parameters zouden wel eens gevoeliger kunnen zijn dan expiratoire parameters zoals de FEV1, omdat de luchtwegcompressie bij de expiratie het effect van luchtwegverwijders zou kunnen maskeren. In de literatuur vinden we steun voor deze gedachte bij O'Donnell en Taube [7,8]. Zij vonden na het geven van bronchusverwijders, een veel sterkere overeenstemming, tussen de verbetering in het gevoel van benauwdheid en de verbetering gevonden bij de inspiratoire longfunctie parameters (ILP) en met name het geforceerde inspiratoire volume in de eerste seconde na een volledige uitademing (FIV1) en de longcapaciteit na een normale uitademing (IC). In ons onderzoek trachten we de vraag te beantwoorden of inademingsparameters beter zijn dan de FEV1 om bronchusverwijding op de sporen.

Onderzoek naar de inademings methode:
Inspiratoire longfunctieparameters kunnen we op twee manieren meten: De eerste manier als onderdeel van de geforceerde luchtstroommeting waarbij de patiënt voorafgaand aan de meting rustig diep inademt en hierna achtereenvolgens geforceerd uitademt en inademt, dus een inademing na een geforceerde uitademing (Fe-methode = Forced expiration method) De tweede manier: we kunnen dezelfde ILP's meten na een rustige uitademing (Se-methode = Slow expiration method). Beide methoden worden in de praktijk gebruikt, maar welke is de beste?
In hoofdstuk 2 staat ons onderzoek naar de inspiratiemethode beschreven [9]. Wij vonden dat de variabiliteit (het omgekeerde van stabiliteit) van de meting van de FIV1 het grootst was bij de Fe-methode en bovendien dat de Se-methode gemiddeld tot een hogere waarde leidde (verschil 0,21 liter). Tot slot vonden we dat het verschil tussen beide methoden afhankelijk was van de ernst van de COPD ingedeeld volgens de z.g. GOLD klasse [1], waarbij het verschil groter werd bij ernstiger vormen van COPD. Hoewel wij een uitstekende correlatie tussen beide methoden zagen, vonden we een te grote variatie tussen metingen binnen het individu. Daarom mogen beide methoden in de praktijk bij een patiënt zeker niet door elkaar gebruikt worden. Wij komen dus tot de conclusie dat we een voorkeur hebben voor

de Se-methode vanwege de betere stabiliteit en het niet afhankelijk zijn van de mate van obstructie bij de COPD patiënt.
Nadat we de beste methode hadden gekozen, onderzochten we hoe vaak we de meting moeten herhalen om een redelijk betrouwbare (reproduceerbare) meting te krijgen en wij vonden dat 5 metingen per patiënt leidden tot een goed compromis tussen het aantal uit te voeren metingen en de maximaal bereikte FIV1 [9]. De mate van ernst (GOLD klasse) was niet van invloed op het aantal uit te voeren metingen.

Onderzoek naar de natuurlijke spreiding van de FEV1
Behalve het belang van de FEV1-meting om de ernst van COPD vast te stellen, wordt de FEV1 ook gebruikt om de mate van gevoeligheid van de luchtwegen te meten na het toedienen van prikkelende stoffen en om de mate van verbetering (reversibiliteit) te meten na een interventie zoals het toedienen van bronchusverwijdende medicijnen. Uit onderzoek bij groepen patiënten met COPD is gebleken dat er een statistisch significante verbetering optrad na het geven van bronchus verwijdende medicijnen (de z.g. sympaticomimetica en parasympaticolytica). Bij de individuele patiënt is het echter moeilijk een significante verbetering te bewijzen en dat geldt zeker voor de patiënt met een ernstig COPD. Indien we daarbij de aanbevelingen voor verbetering van de gezamenlijke Europese (ERS) en Amerikaanse wetenschappelijke "longartsen" vereniging (ATS) zouden opvolgen [10] voldoet slechts 11 % van de patiënten met ernstig COPD aan de criteria genoemd in deze aanbevelingen [11,12]. Deze aanbevelingen zijn: de FEV1 moet minimaal 12% verbeteren ten opzichte van de uitgangswaarde én met een minimum van 200 cc [10]. Met name dat laatste criterium wordt vaak niet gehaald bij patiënten met ernstig COPD. Longartsen zien veel patiënten met ernstig COPD.
In hoofdstuk 3 van het proefschrift wordt het onderzoek naar de natuurlijke spreiding van de FEV1 gepresenteerd. Wij onderzochten de literatuur genoemd in deze aanbeveling en die daarna is verschenen en vonden dat geen van de onderzoekers de door Bland en Altman [4] geadviseerde aanpak heeft gebruikt, terwijl deze methode van aanpak zelfs door de auteurs van de ATS-ERS taskforce aanbevolen werd [10]. Wij onderzochten bij patiënten met COPD de natuurlijke spreiding (RV=random variation) van de FEV1-parameter en onderzochten niet alleen de mate van spreiding (CR=coëfficiënt of repeatability). Daarnaast onderzochten we of deze spreiding afhankelijk

was van de uitgangswaarde van de FEV1. We gebruikten hiervoor een methode die in de praktijk vaak wordt gebruikt om verbeteringen na een interventie uit te drukken, te weten: als absolute verbetering b.v. 200 ml, of als procentuele verbetering van de uitgangswaarde b.v. 12% of als procentuele verbetering ten opzichte van de voorspelde waarde b.v. 9%. Het type spreiding dat voor alle uitgangswaarden gelijk is noemen we homoscedastisch, het type dat afhankelijk is van de uitgangswaarde noemen we heteroscedastisch.
Het kennen van de natuurlijke spreiding is belangrijk omdat men bij een individuele patiënt alleen kan aantonen dat een medicament werkt, indien de verandering die optreedt na het toedienen ervan de natuurlijke spreiding (RV) overstijgt. Bij een homoscedastische spreiding hoeven we maar 1 getal te kennen voor alle uitgangswaarden van de FEV1 b.v. 12%, maar bij een heteroscedastische spreiding is dit in feite onmogelijk, omdat bij iedere uitgangswaarde een andere spreidingsmaat van toepassing is. In een steekproef van 79 patiënten met COPD werd de FEV1 5 x binnen een dag gemeten. Het verschil in FEV1-waarden genomen met een uur tijdsverschil, werd berekend en de spreiding daarvan werd bepaald. We vonden dat het type spreiding homoscedastisch is indien we het verschil tussen de metingen uitdrukten als het percentage van de uitgangswaarde maar ook van het percentage van de voorspelde waarde. De spreiding was echter heteroscedastisch indien we het verschil uitdrukten in het verschil in liters. We vonden voor de coëfficiënt van herhaalbaarheid (CR) dat 12 % een goede waarde is gezien vanuit de uitgangswaarde. Echter het criterium van 200 cc is niet zo waardevol vanwege de heteroscedastische spreiding, we kunnen in dat geval dus niet volstaan met één getal. De consequentie is dat indien we het 200 cc-criterium negeren bij patiënten met een ernstig COPD, er veel vaker een statistisch significante verbetering kan worden gevonden na het toedienen van bronchusverwijders, dan mogelijk is met de aanbevelingen van de ERS en ATS.

Onderzoek naar de natuurlijke spreiding van de inspiratoire parameters.
In hoofdstuk 2 beschreven we al hoe we inspiratoire parameters het beste konden meten, in hoofdstuk 4 gaan we in op de natuurlijke spreiding (zowel het type als de mate) van de inspiratoire longfunctieparameters bij dezelfde groep patiënten als beschreven bij de FEV1. De volgende parameters

werden gemeten: het geforceerde inspiratory volume in 1 second (FIV1), inspiratory capacity (IC), maximal inspiratory flow at 50% (MIF50) en peak inspiratory flow (PIF). Ook hierbij vonden we voor al deze parameters een heteroscedastische spreiding indien we verschillen uitdrukten in liters of liters per seconde. Een homoscedastische spreiding vonden we indien we uitgingen van de uitgangswaarde. Dus stellen we voor om net als bij de FEV1, de verschillen uit te drukken in percentage van de uitgangswaarde (een percentage van normaal is niet te geven omdat er geen geaccepteerde normaalwaarden beschikbaar zijn). Wij vonden voor de natuurlijke spreiding de volgende waarden: IC, 19%; FIV1, 14%; PIF, 18%; MEF50, 21%.

Effect van Pursed-lips breathing op ILP en FEV1.

Nu we de natuurlijke spreiding van de FEV1 en de inspiratoire parameters kennen, is het mogelijk deze parameters voor en na verschillende interventies te meten zoals het toedienen van kort werkende luchtwegverwijders, histamine provocatie en het met geperste lippen uitademen (PLB=pursed lips breathing). Dat laatste beschrijven we in hoofdstuk 5 [13] en de kortwerkende bronchusverwijders in hoofdstuk 6 [14].

Bij patiënten met COPD verbetert PLB de gasuitwisseling in de longen en de hyperinflatie van de longen gemeten via een methode die met elektro-optische koppeling werkt. Het effect van PLB op inspiratoire parameters was echter niet bekend. Wij maten dit effect in een groep van 35 COPD patiënten met een ernstige luchtwegobstructie (FEV1< 50% van de voorspelde waarde). Wij maten vooraf, onmiddellijk na PLB, en nog eens 5 minuten later de volgende primaire parameters: forced inspiratory vital capacity (FIVC), inspiratory capacity (IC), forced inspiratory volume in first second (FIV1), maximal inspiratory flow at 50% of VC, en peak inspiratory flow (PIF). Patiënten werden ook getest op de volgende secundaire parameters:vital capacity (VC), FEV1, adem frequentie (BF), CO2 spanning aan het einde van een normale uitademing (ET-CO2) en de zuurstofsaturatie (SO2). SO2 is een maat voor het zuurstof gehalte gemeten met een pulseoximeter.

Van de primaire parameters verbeterde alleen de IC significant; van de secundaire parameters verbeterde de gemiddelde SO2 met 1% ($p = 0.005$) en de gemiddelde ET-CO2 en BF met 3.2 mm Hg en 3.1 ademteugen per minuut. Na 5 minuten waren deze effecten duidelijk minder.

Effect van kortwerkende bronchusverwijders op ILP en FEV1[14].
Het doel van deze studie was het effect te meten van salbutamol, ipratropium en placebo op de ILP en de FEV1 en dit effect vervolgens te vergelijken met veranderingen in het subjectieve gevoel van benauwdheid, gemeten op een visueel analoge schaal (VAS). Alle gemiddelde waarden van de ILP en de FEV1 verbeterden significant na toediening van de medicijnen (met uitzondering van de FIF50 voor ipratropium) maar niet na toediening van de placebo. Het benauwdheidsgevoel verbeterde ook bij placebo toediening significant maar we vonden in tegenstelling tot Taube et al , geen significante overeenstemming tussen de VAS en de verbetering van de ILP na bronchusverwijders. Verder vonden we geen significante verschillen in het aantal patiënten dat meer dan de natuurlijke spreiding verbeterde (responders) op de inspiratoire parameters en op de FEV1. Getalsmatig vonden we zelfs meer responders met de FEV1 .

Doelmatigheidsonderzoek.
Astma en COPD komen vaak voor en longfunctieonderzoek is noodzakelijk om de diagnose te bevestigen. Daarom is het belangrijk, voor het gemak van de patiënt maar ook uit kosten-overwegingen, dat dit onderzoek zo effectief en efficiënt mogelijk wordt uitgevoerd.
Bezoeken op de polikliniek en longfunctietesten (LFT) vormen een belangrijk deel van de financiële lasten. Om effectiever en efficiënter te werken stelden wij een longfunctieprotocol op voor patiënten met een obstructieve longaandoening. Echter niet eerder was onderzocht of en in welke mate dit tot kostenbeheersing zou kunnen leiden.
In hoofstuk 7 doen wij verslag van ons doelmatigheidsonderzoek. Een prospectief, enkel blind en gerandomiseerd onderzoek werd uitgevoerd. In de controle groep (CG) werden alle LFT die door de specialist waren aangekruist op de longfunctie-afdeling uitgevoerd.
In de experimentele groep (EG) werden longfunctietesten uitgevoerd volgens het door ons opgestelde longfunctieprotocol. Het doel was de kosten die nodig waren om tot de uiteindelijke diagnose te komen te vergelijken. In totaal namen 179 patiënten met verdenking op een obstructief lijden deel aan de studie: 86 in de CG and 93 in de EG. Het gemiddelde aantal testen om tot een diagnose te komen was 3,8 in de CG versus 2,9 in the EG ($P<0,001$). Het gemiddelde aantal overbodige LFT voordat de diagnose gesteld kon worden was 1.2 in de CG versus 0.08 in de EG. Het aantal

patiënten dat extra op de poli terug moest komen omdat de diagnose niet rond was, kwam hoger uit in de CG dan in de EG. De totale kosten per gestelde diagnose (van poli bezoeken en LFT) bedroegen €227 in de CG versus €181 in de EG.
De conclusie was dat LFT volgens ons longfunctieprotocol doelmatiger zijn dan het aankruisen van LFT door artsen bij patiënten verdacht van een obstructieve longaandoening.

Discussie:
Dit proefschrift voegt volgens ons het volgende toe aan wat we al weten over het basale longfunctieonderzoek bij patiënten met COPD:

- Inspiratoire longfunctieparameters kunnen beter bepaald worden na een rustige uitademing, dan na een geforceerde uitademing.
- Vijf inspiratoire longfunctieoefeningen zijn nodig om een optimaal resultaat te krijgen ten aanzien van betrouwbaarheid en haalbaarheid.
- De natuurlijke spreiding van de door ons onderzochte inspiratoire parameters en de FEV1 was ongewenst heteroscedastisch wanneer wij de verandering uitdrukten in liters of liters per seconde maar veel meer gewenst homoscedastisch wanneer we deze uitdrukten als percentage van de uitgangswaarde.
- We vonden geen aanwijzingen voor de validiteit van het ERS-ATS criterium van minimale verbetering van tenminste 200 cc voor de FEV1 na bronchusverwijders om van een significante verbetering te mogen spreken.
- De inspiratoire parameters doen het niet beter dan de FEV1 om een significante verbetering aan te tonen (meer verbetering dan de natuurlijke spreiding) na het toedienen van kortwerkende luchtwegverwijders.
- De Pursed-lips adem-techniek resulteert alleen in een significante verbetering van de inspiratoire capaciteit (een maat voor hyperinflatie) maar niet van de longfunctie parameters die stroomsnelheid meten.
- Bij patiënten verdacht voor een obstructieve longaandoening is het doelmatiger te werken met een gedefinieerd longfunctieprotocol dan dat artsen naar eigen inzicht los longfunctie-onderzoeken aankruisen.

Op de vraag of inspiratoire longfunctieparameters gemeten bij patiënten met COPD veel toevoegen aan het routine longfunctieonderzoek waarbij vooral de FEV1 en FVC worden gemeten moet het antwoord ontkennend zijn. Wij concluderen dat voor routine longfunctieonderzoek er geen harde noodzaak bestaat om inspiratoire longfuncties te meten want met de FEV1 vinden we minimaal evenveel patiënten die significant verbeteren op voorwaarde dat we alleen het criterium gebruiken van 12% verbetering t.o.v. de uitgangswaarde en het criterium van tenminste 200 cc laten varen. Dit was niet wat wij verwachtten of hoopten bij de aanvang van onze studies. Wij verwachtten dat inspiratoire longfunctieparameters het wel eens veel beter zouden kunnen doen. Een reden dat Taube et al. tot andere bevindingen kwamen is mogelijk de patiënten selectie.
Wij onderzochten meer patiënten met minder ernstig COPD, terwijl Taube een steekproef nam bij patiënten met een lager gemiddelde FEV1 t.o.v. de voorspelde waarde dan in onze steekproef.
Een tweede reden is dat wij alleen patiënten onderzochten met een stabiel COPD, dus die al enige tijd geen prednison of antibiotica nodig hadden. Een derde reden is dat veel patiënten met stabiel COPD geen kortademigheid in rust ervaren. Waarom zouden zij zich dan minder benauwd voelen na het geven van bronchusverwijders? Misschien was er wel verschil merkbaar geweest, als we keken naar kortademigheid bij inspanning maar dat hebben zowel Taube et al.en wij niet onderzocht. Tot slot vonden we dat bij een deel van onze COPD patiënten de vitale capaciteit al binnen de seconde was bereikt en dat maakt de FIV1 minder gevoelig om veranderingen te kunnen meten. Hoe Taube hier mee omging heeft hij niet beschreven. Het zou kunnen dat we dat probleem hadden kunnen ondervangen wanneer we niet van de FIV1 waren uitgegaan maar van de FIV in een halve seconde FIV1/2 .

Verder onderzoek:
De waarde van inspiratoire parameters bij verschillende interventies wordt door ons nog onderzocht, zoals bij toedienen van histamine, lang werkende bronchusverwijders en van inhalatiesteroiden of orale steroïden.
Gedurende de looptijd van onze studie werden meer vragen opgeroepen:

- Kunnen we in een aantal gevallen in plaats van de nogal belastende histamine provocatie test ook volstaan met de NO meting in de uitademingslucht en deze meting integreren in ons longfunctieprotocol?

- Kunnen we om de reversibiliteitsmeting efficiënter te maken in één keer zowel de reactie op salbutamol testen als de daarna toegevoegde verbetering door middel van ipratropium?
- Kunnen we om de reversibiliteitstest nog efficiënter te maken in een keer zowel de salbutamol als de ipratropium toedienen en hun relatieve werking testen na respectievelijk 15 en 45 minuten?
- Wat is de natuurlijke spreiding voor en na maximale vrijwillege ventilatie van de inspiratoire capaciteit?
- Indien we de hierboven beschreven spreiding weten, kunnen we onderzoeken of deze test voor dynamische hyperinflatie gevoeliger is dan de FEV1 om responders te vinden zonder de noodzaak hiervoor inspanningstesten te doen.

Referenties

1. From the *Global Strategy for the Diagnosis, Management and Prevention of COPD*, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2009. Available from: http://www.goldcopd.org.

2. From the Global Strategy for Asthma Mangement and Prevention, Global initiative for Asthma (GINA). 2010. Available from http://www.ginasthma.com.

3. Archibald CJ, Guidotti TL: Degree of objectively measured impairment and perceived shortness of breath with activities of daily living in patients with chronic obstructive pulmonary disease. *Can J Rehab* 1987, 1:45.

4. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986, 1:307-310.

5. Cazzola M: Should the choice of a long-acting bronchodilator in the long-term therapy of chronic obstructive pulmonary disease depend entirely on the onset of action? *Respiration* 2006, 73:410-411.

6. Guyatt GH, Townsend M, Nogradi S, Pugsley SO, Keller JL, Newhouse MT: Acute response to bronchodilator. An imperfect guide for bronchodilator therapy in chronic airflow limitation. *Arch Intern Med* 1988, 148:1949-1952.

7. Taube C, Lehnigk B, Paasch K, Kirsten DK, Jorres RA, Magnussen H: Factor analysis of changes in dyspnea and lung function parameters after bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000, 162:216-220.

8. O'Donnell DE, Forkert L, Webb KA: Evaluation of bronchodilator responses in patients with "irreversible" emphysema. *Eur Respir J* 2001, 18:914-920.

9. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Recommendations for the measurement of FIV(1) values in chronic obstructive pulmonary disease. *Respiration* 2008, 76:46-52.

10. Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J et al.: Interpretative strategies for lung function tests. *Eur Respir J* 2005, 26:948-968.

11. Han MK, Wise R, Mumford J, Sciurba F, Criner GJ, Curtis JL, Murray S, Sternberg A, Weinman G, Kazerooni E et al.: Prevalence and clinical correlates of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 35:1048-1056.

12. Hansen JE: Assessment of bronchoreversibility in severe emphysema. *Eur Respir J* 2010, 36:972.

13. Visser FJ, Ramlal S, Dekhuijzen PN, Heijdra YF: Pursed-lips breathing improves inspiratory capacity in chronic obstructive pulmonary disease. *Respiration* 2011, 81:372-378.

14. Ramlal SK, Visser FJ, Hop WC, Dekhuijzen PN, Heijdra YF: Reversibility of inspiratory lung function parameters after short-term bronchodilators in COPD. *Respir Physiol Neurobiol* 2010, 173:58-63.

10

Dankwoord en Curriculum Vitae

Dankwoord

Een onderzoek bedenken, uitvoeren en schrijven doe je met een hele groep mensen. Die staan je allemaal bij in het proces naar dit proefschrift, maar krijgen daar slechts beperkt eer voor. Daarom vind ik het belangrijk om deze mensen te noemen en te bedanken.
In de eerste plaats natuurlijk dank aan de patiënten. Zij hebben vaak hun vrije tijd ingeleverd en regelmatig hun reguliere medicatie moeten uitstellen om dit onderzoek mogelijk te maken. Zonder onze patiënten was dit onderzoek onmogelijk geweest.
Dank aan mijn promotor prof dr. P. N. Richard Dekhuijzen en copromotor dr. Yvonne F. Heijdra. Zonder jullie aanbevelingen, brainstorm- en evaluatiesessies, was dit niet tot een goed einde gekomen. Met name Yvonne heeft heel veel werk verzet om mijn soms onbegrijpelijke uiteenzettingen over het homo- of heteroscedastisch zijn van spreidingen, op een verantwoorde wijze aan het papier toe te vertrouwen. Sunil, mijn belangrijke medeonderzoeker en (mede) auteur van vele artikelen: aan jou veel eer. Zonder jouw aanzet en wens om te promoveren samen met mij (jouw proefschrift komt hoop ik ook snel), was ik hier mogelijk niet meer aan begonnen. Jij zorgde ook dat veel patiënten enthousiast werden om mee te doen en, eerlijk gezegd, als ik jouw patiënt was geweest, had ik me ook laten verleiden. Door jouw voortdurende interacties met de commissie mensgebonden onderzoek werden alle interventiestudies toegestaan. Mijn andere mede auteur Milena van der Vegt: veel werk verrichtte je voor het doelmatigheidsonderzoek. Op heel wat woensdagmiddagen hebben we data uit de statussen gehaald en ze naar tabellen omgezet. Ook heb jij een poster over dit onderzoek gepresenteerd en heb je intensief meegeschreven aan het uiteindelijk artikel. Ben Pelzer en Gert Jan van der Wilt, dank voor de ondersteuning betreffende methodiek en statistiek. Julius Jansen, jij zette ons aan tot het doelmatigheidsonderzoek en sprak met Richard en Yvonne over de ambitie een promotieonderzoek van het inspiratieproject te maken voor zowel Sunil als mij.
Zonder metingen geen onderzoek, en dus ben ik dank verschuldigd aan Bard Kelder en Renate van Uden, die meedachten en soms onhaalbare hersenspinsels van Sunil en mij corrigeerden. Niet minder dank aan Nicole Gasperz, Stefanie Engelen en Katja van Hulsdonk; zij namen het gros van de metingen voor hun rekening.
Anneloes Niesink en Michiel Spanbroek hielpen mee met de dataverzameling

en schreven voor hun wetenschappelijke stage een verslag. Ook jullie hebben bijgedragen. Michiel, jij veel succes toegewenst met jouw desmosineonderzoek.
Ina Mulder (wetenschapsbureau CWZ) en Rigoni de Abreu wil ik bedanken voor het nakijken van de inleiding en samenvatting in respectievelijk het Engels en Nederlands en Karien Pittie voor het dankwoord en curriculum vitae.
GlaxoSmithKline jullie hielpen ons zonder restricties met een bijdrage aan ons researchfonds.
Mijn maatschap heeft mij gesteund. Yvonne Berk nam na 2009 mijn taak als medisch manager over en ik heb weinig nieuwe taken daarvoor in de plaats gekregen. Willem van den Berg en Rob Janssen, jullie namen mijn logistieke taken over, met name de financiële taken in de maatschap en de medische staf. René Termeer, op de achtergrond heb je veel werk verzet als voorzitter van de maatschap en ik heb altijd goed met je kunnen overleggen. Anneke van Veen, jij nam de automatisering en een deel van de longfunctiewerkzaamheden van mij over.
Het thuisfront wil ik zeker niet vergeten. De laatste jaren waren zeer bewogen door het overlijden van onze ouders, Elines oppas Louk van der Weerden, goede vriend Wim Westerhof en nichtje Eva.
Mijn vrouw Marian en onze kinderen Fleur, Michiel en Eline en hun partners Wesley en Hedwich hebben me misschien wat te vaak achter de PC zien zitten. Eline, dank voor de mooie omslag van dit boekje; een foto die je maakte in Tenero aan het Lago Maggiore. Marian jij hebt mij altijd gesteund en dus ook weer in de afgelopen jaren en zeer geholpen bij de voorbereidingen voor het promotie feest. Ik ben trots op jou en onze kinderen.
Mijn paranimfen dochter Fleur en vriend Jos Schulte, dank voor jullie praktische steun en het bijstaan tijdens de feestelijkheden.
Tot slot was het voor mij een zeer leerzame zoektocht in het land van de wetenschap. Ik heb het nodige over statistiek en het toepassen daarvan geleerd. Het meest inspirerend vond ik de data- analyse en het meest taaie, het opschrijven van de verworven kennis. Maar gemiddeld heb ik ervan genoten en is een oude droom, eens te promoveren, uitgekomen.

Malden, februari 2012

Frank Visser

Curriculum vitae

Frank Visser is geboren op 26-november 1950 in Terneuzen, en in 1982 gehuwd met Marian Wullems. Ze hebben 3 kinderen: Fleur 1983, Michiel 1985 en Eline 1994. Hij studeerde Geneeskunde in Utrecht van 1970-1977 en werkte daar ook als student-assistent op de afdeling fysiologie van 1972 tot 1974.
De opleiding Interne geneeskunde volgde hij van 1977-1979 in het Elisabeth ziekenhuis in Tilburg. Tijdens een stage longziekten bij Jan Molkenboer is de interesse voor longziekten en tuberculose bij hem gewekt. Hij volgde de opleiding tot longarts van 1979-1982 in het Catharina ziekenhuis te Eindhoven (opleider dr. C.A.F. Jansveld). In Eindhoven ontwikkelde hij een computerprogramma om longfuncties te verslaan en te interpreteren. Na zijn opleiding werd hij longarts in het Canisius Wilhelmina Ziekenhuis (CWZ) in Nijmegen waar hij nog steeds werkt. Tot in 2011 werd gebruik gemaakt van het al eerder genoemde interpretatieprogramma voor longfunctieonderzoek Naast zijn werk als algemeen longarts met als aandachtsgebied longfunctie richtte hij in 1993 een slaaplaboratorium op. Inmiddels zijn ademhalingsstoornissen tijdens de slaap een belangrijk aandachtsgebied geworden binnen het CWZ.
Bestuurservaring deed hij op in het CWZ als lid van het stafbestuur en later als medisch manager van de longafdeling. Binnen de Nederlandse Vereniging van Artsen voor Longziekten en Tuberculose (NVALT) was hij voorzitter van de Commissie voor classificatie van longziekten. Deze classificatie samen met een informatiemodel voor een elektronisch patiënten dossier werd gebruikt voor een model elektronisch medisch dossier binnen de Orde van medisch specialisten. Een computer programma dat hij naar aanleiding van dit bovenstaande model ontwikkelde, is nu nog in gebruik als elektronisch medisch dossier binnen de afdeling longziekten van het CWZ. In 2011 werd de afdeling longziekten aangesloten bij het ziekenhuis EPD waarin functionaliteit van het ‘’eigen programma’’ werd ingebouwd.
Daarnaast was Frank Visser gedurende zijn opleiding en later van 2002 tot 2010 lid van het Consilium van de NVALT in een tijd dat de opleidingen en de wijze van het visiteren van opleidingen flink aan wijzigingen onderhevig waren. Nu is hij nog lid van de Commissie voor geschillen van de KNMG.
Tijd voor hobby’s was en is er gelukkig ook:
Tennis, skiën, schaatsen, beleggen en zingen in het koor Malle Muze zijn hobby’s die hij beoefent.