OCCASIONAL REVIEW

Assessment of progression of COPD: report of a workshop held in Leuven, 11–12 March 2004

M Decramer, R Gosselink, M Rutten-Van Mölken, J Buffels, O Van Schayck, P-A Gevenois, R Pellegrino, E Derom, W De Backer

Thorax 2005;60:335-342. doi: 10.1136/thx.2004.028712

Recently performed long term trials have enhanced the insight into the assessment of progression of COPD. The present review focuses on the initial assessment of COPD in general practice and the assessment of disease progression. Several variables may be used to assess this progression, all of which are associated with significant methodological problems. Finding the appropriate mix of outcome measures to capture all aspects of disease progression is a significant challenge.

uring the last decade insight into the natural course of chronic obstructive pulmonary disease (COPD) has grown considerably. This insight largely stems from four large scale and long term studies on the effects of inhaled corticosteroids,1-4 large scale 1 year studies with tiotropium,5 6 and with the combination of long acting β2 agonists and inhaled corticosteroids.7 8 These studies have unravelled a number of methodological problems associated with large scale long term studies and have enhanced insight into the phenotype. In addition, other studies have enhanced our knowledge of how to diagnose the disease in the early stages.9-11 The diagnosis of COPD has also been modified to a significant extent by the recent GOLD guidelines.¹² Although, classically, COPD has been defined as a disease associated with largely irreversible airflow obstruction, it has become increasingly clear that up to 40% of patients with COPD show significant reversibility, particularly in terms of vital capacity or inspiratory capacity.13-15 It therefore appears more appropriate to define COPD as "not fully reversible airflow obstruction" rather than "irreversible airflow obstruction", as in the GOLD definition.

At the same time, new ways to assess the progression of COPD besides the classical annual decline in forced expiratory volume in 1 second (FEV₁) have become available.^{16–18} These new methods may facilitate future long term studies although this is not yet certain. Moreover, significant methodological problems with long term studies have become apparent. Firstly, in most long term studies drop out is important, ranging from 25% to 53% over 3 years. This not only causes important problems regarding the power required to demonstrate differences between treatment groups, but also creates a significant problem for the analysis of the results

as the population at the end of the trial is no longer representative of the population at the beginning. Furthermore, drop out in the major COPD trials is not completely random, but is at random or informative, with the worst patients dropping out earlier than the better ones.¹⁹ Drop out is also greater in the placebo group than in the active treatment group, and patients dropping out from the active treatment group are usually better than those dropping out of the placebo group. The combination of these factors poses an important methodological problem and calls for more sophisticated statistical models.

The purpose of the present review is to address the challenges of assessing the natural course of COPD from the perspective of what is known from the large long term trials conducted to date. We will first address the current view on the assessment of incipient COPD and then consider the assessment of the progression of COPD.

INITIAL ASSESSMENT IN GENERAL PRACTICE

It is clear that COPD cannot be diagnosed on the basis of symptoms alone. In general, symptoms correlate poorly with pulmonary function. Although a large fraction of patients with COPD have symptoms of cough, sputum and/or dyspnoea, the diagnosis requires demonstration of not fully reversible airflow obstruction and, hence, requires spirometric tests.¹²

COPD has been defined in different ways. The European Respiratory Society (ERS) definition states that COPD is characterised by airflow obstruction that is irreversible.20 The more recent GOLD guidelines¹² define COPD as a disease with airflow obstruction that is not fully reversible. The latter is presently considered as the most adequate. Airflow obstruction should always be documented with spirometric tests showing a decrease in the FEV₁/FVC ratio below the 5th percentile21 or below 70%.12 Whether this latter definition of airflow obstruction is also adequate for an elderly population remains a matter of debate.10 In a few patients, however, airflow obstruction may occur with a decrease in FEV1 of a similar extent to vital capacity (VC) so that their ratio remains normal.21 Under these conditions, a normal total lung capacity (TLC) is the

Abbreviations: COPD, chronic obstructive pulmonary disease; FEV $_1$, forced expiratory volume in 1 second; FVC, forced vital capacity; MCID, minimal clinically important difference; 6MWD, 6 minute walking distance; QoL, quality of life; $\Delta Vo_2 max$, change in maximal oxygen consumption; $\Delta Wmax$, change in maximal work load

See end of article for authors' affiliations

Correspondence to: Professor M Decramer, Respiratory Division, University Hospital, Herestraat 49, 3000 Leuven, Belgium; Marc. Decramer@uz.kuleuven. ac.be

Received 18 May 2004 Accepted 18 December 2004 only parameter capable of differentiating obstruction from restriction. Moreover, a large fraction of patients do not report or clearly underreport symptoms.²² Most patients in the earlier stages of GOLD I and II are likely to be treated in general practice rather than in hospital.

Is the GP able to establish this diagnosis? We found no clearcut answer to this question in the literature. While there is little doubt about the role of spirometric testing in the GP surgery in detecting COPD, there is little evidence that it is sufficiently accurate for diagnostic and staging purposes. One study in Australia showed that GPs performed poorly in spirometric testing even after instruction.²³ Indeed, only 35% of the tracings generated by GPs who attended an introductory workshop on spirometry reached the American Thoracic Society (ATS) acceptability criteria; for those who did not attend the workshop, only 16% of the tracings they generated reached these acceptability criteria. The major problem appeared to be failure to reach the end of test criteria. This was in line with other studies performed in general practice.24 25 More recent studies, however, have shown a substantially better performance by GPs.9 26 The latter study reported an equal amount of non-reproducible tests (16-18%) in general practice as in a regular pulmonary function laboratory. This may be due in part to the involvement of practice assistants trained with a 5 hour course in spirometric testing. More research is needed on the role of these practice assistants in improving the performance of spirometric tests in general practice.

A recent study advocated the use of peak flow rate measurements in the diagnosis of COPD.27 This study showed that a peak flow rate of less than 80% had a sensitivity of 91% and a specificity of 82% in detecting COPD. This sensitivity was even greater for patients with moderate to severe disease. However, with a positive predictive value of 30%, peak flow measurement is clearly insufficient for diagnostic purposes. In addition, the study was a retrospective subgroup analysis which may limit its value. A complete flow-volume loop is an important diagnostic element in the diagnosis of COPD because incipient obstruction is primarily localised in the small airways and does not markedly affect the peak expiratory flow rate (see above).28 One study suggested that patients with mild to moderate COPD may be reluctant to be referred to a pulmonologist. In this study GPs performed 3209 spirometric tests in patients at risk and found 723 new cases of probable COPD (23%). Only 278 of these subjects (39%) agreed to consult a pulmonologist.29 Whether this is a generalised problem remains to be studied.

ASSESSMENT OF DISEASE PROGRESSION OF COPD Decline in FEV₁

The annual decline in FEV $_1$ has been the standard way of assessing the progression of COPD for several decades. ³⁰ In normal subjects the annual decline amounts to 20–30 ml/year compared with an average of 60 ml/year in patients with COPD. ³¹ In the 1990s several randomised controlled studies on the effects of inhaled steroids were conducted with the annual decline in FEV $_1$ as the primary outcome variable, ^{1–4} largely building on the design of the Lung Health Study I. ³¹ In these studies the annual decline ranged from 42 ml to 69 ml (table 1).

The problems with measuring the annual decline in FEV₁ are essentially threefold. Firstly, the variability in the annual decline is large which has been known ever since the first publication.³⁰ In the published studies the SD of the decline ranges from 58 ml/year in the Lung Health Study⁴ to over 75 ml/year in ISOLDE³ and to 100 ml in EUROSCOP.¹ Given that published studies assume arbitrarily that a difference in annual decline of 20 ml/year^{1–3} would be clinically significant, the sample size of the patients completing the study would

Table 1 Annual decline in FEV₁ in ml/year versus mean (SD) FEV₁ in ml: the accuracy of the mean FEV₁ measurement is shown in square brackets

Study	Active	Placebo
EUROSCOP1	2530 (640) ml	2540 (640) ml
	57 ml/year [127 ml]	69 ml/year [127 ml]
Copenhagen	2360 (790) ml	2390 (860) ml
City Lung Study ²	45 ml/year [118 ml]	42 ml/year [120 ml]
ISOLDE ³	1420 (470) ml	1400 (480) ml
	50 ml/year [71 ml]	59 ml/year [70 ml]
LHS II⁴	2280 (620) ml	2220 (650) ml
	44 ml/year [114 ml]	47 ml/year [111 ml]
BRONCUS ^{33 51}	1600 (380) ml	1650 (390) ml
	56 ml/year [158 ml]	47 ml/year [160 ml]

need to range from at least 200–800 for 80% power and a significance level of 5%, which are minimal requirements. The standard deviations given pertain to at least 3 years follow up and four measurements per year, which appears to be optimal.³²

Secondly, due to the limits of the accuracy and precision by which FEV₁ can be measured, a follow up of at least 3 years is required to estimate the decline in FEV1. Indeed, if we assume a mean FEV₁ of 1.60 l, then—given an accuracy of 5%—the minimal volume that can be measured accurately would be 80 ml. Table 1 shows the annual decline and the mean (SD) FEV₁ measured in the major studies. As can be seen, the annual decline is clearly below the accuracy of the mean FEV₁ measurement in these trials. Given an annual decline of 30-60 ml/year, clearly more than 2 years is needed to have differences that exceed the accuracy limits of spirometry. Because of this long term follow up, drop out of patients will be important, ranging from 25% to 50%. Table 2 shows the drop out rates in the placebo and active treatment groups in recently published major studies. Because of the high requirements for accuracy and precision, concerted efforts were made in most trials to reduce variability of spirometric testing as much as possible. A number of measures were used to this end-for example, limiting the number of centres;31 training sessions for pulmonary function technicians;1 3 33 use of an identical spirometer in all centres;13 measurements always at the same time of the day to avoid diurnal variation;1 33 measurements after administration of a standardised dose of bronchodilators.

Thirdly, the drop out rate in long term studies is usually large (table 2) and is not completely at random. 19 Indeed, the worst patients drop out first so that the population at the end of the trial is very different from the intention to treat population. Since baseline characteristics of future drop outs are usually different from those who complete the study, it is possible to correct for this differential drop out by using mixed effect models taking the baseline characteristics as covariates. If, however, this drop out is informative, 19 34 then correction is not really possible. Informative drop out is defined as drop out that is not predictable from other measurements observed but depends on the unobserved measurements—that is, those measurements that would have been observed if the patient would not have dropped out from the study. Since in most studies conducted so far drop out is greater in the placebo group than in the active treatment group (table 2), it is also important to use all available data from those completing the study and from drop outs in the analysis and to use a mode of analysis that takes this factor into account. When this is taken into account, the above calculated sample size would need to range between 1000 and 1500 patients at least. This means

Study	Duration (years)	Reference	Active	Placebo	No randomised
EUROSCOP	3	Pauwels et al	27%	30%	1277
ISOLDE	3	Burge et al ³	43%	53%	<i>75</i> 1
Copenhagen City Lung Study	3	Vestbo et al ²	25%	35%	290
LHŚ II	3	LHS Group⁴	Not reported	Not reported	1116
TRISTAN	1	Calverley et al ⁸	29%	38% '	1465
Symbicort	1	Szafranski <i>et al</i>	30%	44%	812
,	1	Calverley et al ⁸	37%	41%	1022
Tiotropium	1	Vincken <i>et al</i> ⁵	15%	21%	535
	1	Casaburi et al ⁶	19%	28%	921
BRONCUS	3	Decramer et al ³³	27%	37%	573
Average			27%	37%	

that, for most long term studies, the sample size is just at or below this limit (table 2). It should be noted that such a study would then still only be minimally powered for the primary outcome variable. If subgroup analysis is undertaken, substantially more patients would be needed.

Although in the last decades measurement of the decline in FEV_1 has been the standard measure of disease progression of COPD, there is little supportive evidence to show that it is the best measure. Given the limitations associated with the pathophysiology of airflow obstruction outlined recently, ^{13–15} it may reflect more loss of elastic recoil and loss of interdependence between airway and lung parenchyma than airway obstruction. Relatively few data are available on decline in VC, inspiratory capacity (IC), or transfer factor.

Assessment of progression by CT scanning

The use of CT scans to assess the progression of COPD has been limited to date.¹⁸ In the study by Dirksen *et al*¹⁸ the variance in the progression of COPD estimated by CT scanning was considerably smaller than with the decline in FEV₁. This is promising as it would allow studies on the progression of COPD to be performed with considerably fewer patients. We await the confirmation of these results by others

Decline in exercise capacity

Over the past decade the decline in FEV₁ as a function of time has been used as a marker for the natural history of COPD and as an outcome measurement to document the efficacy of therapeutic interventions such as smoking cessation (see above). Although COPD is, by definition, characterised by airflow obstruction, it is now increasingly regarded as a systemic disease.35 The hormonal, metabolic, and musculoskeletal implications of the systemic effects of COPD involving oxidative stress, inflammatory mediators, cytokines, and endocrine hormones result in loss of body weight, osteoporosis, skeletal muscle dysfunction and wasting. One of the markers of the systemic effects of COPD is the decrease in exercise capacity which results from an interaction between decline in pulmonary function and some of the extrapulmonary effects of COPD such as cardiovascular impairment, skeletal muscle weakness, and alteration in body composition. It is increasingly clear that some markers may have important effects on prognosis.36

The decline in exercise tolerance over years has only recently been systematically assessed in patients with COPD.³⁷ In this study, 198 patients with COPD and 41 age matched controls were followed for 2 years. The 6 minute walking distance (6MWD) decreased in the COPD group from 238 (107) m to 218 (112) m. In patients who died within 2 years the mean decline in 6MWD averaged 40

(47) m/year compared with a decline of 21 (31) m/year in those who survived. This study also showed that the 6MWD and its decline appeared to be an independent and possibly a better predictor of survival than FEV₁.³⁷

Data obtained from trials designed to assess the effects of pulmonary rehabilitation programmes or lung volume reduction surgery in COPD confirm that the decline in exercise tolerance is a fact and may be even more pronounced than was previously thought. These are summarised in table 3. The data presented in long term studies should be interpreted with caution. Indeed, the "survival of the fittest" generally tends to underestimate the decline, since patients exhibiting lower initial values or a more rapid decline tend to die at an earlier time point (drop out not completely at random, see above). This might eventually result in an overestimation of "mean" exercise tolerance after 1 or 2 years.

In two studies substantial decreases in walking distance occurred over 12 months³⁸ and 18 months³⁹ in the patients with COPD who were not randomised to participate in a rehabilitation programme. Interestingly, exercise tolerance of rehabilitated patients did not deteriorate in these studies. Surprisingly, exercise tolerance over 24 months did remain stable in the control group of a third long term follow up study.⁴⁰ The reason for this discrepancy is unclear but might have been caused by differences in the degree of severity of COPD or a less liberal use of oral corticosteroids.

Pulmonary rehabilitation does, however, not always "protect" against progressive decline in exercise tolerance. Indeed, a substantial decline has been reported in patients after rehabilitation in two recently published studies. First, control patients included in the NETT trial, who followed a 6-10 week rehabilitation program but did not undergo surgery, exhibited a substantial decline in FEV1 and in maximal work load over the 24 month follow up period.41 Decreases up to 20 Watts or more were observed in 57%, 59%, and 90% of the patients 6, 12, and 24 months after the end of the rehabilitation programme, respectively. These findings are in line with those of Ries et al42 who investigated the effects of a maintenance treatment on the long term effects of an 8 week pulmonary rehabilitation programme in COPD in a controlled study. A significant decline over time was observed in both groups. In the control group 6MWD fell by 44 m and maximal oxygen consumption by 130 ml/min over the 24 month follow up period, whereas FEV₁ decreased by only 70 ml over the same period.42 It thus appears that decline in exercise tolerance is not infrequently encountered in patients with severe COPD (table 3), and that only rehabilitation programmes exceeding 3 months in duration are able to prevent this decline to a certain extent.

Table 3 Decline in exercise tolerance in patients with COPD

No of patients	FEV ₁ (I)	FEV ₁ (%pred)	Follow up (months)	∆6MWD (m)	Δ Shuttle test (m)	∆ Work load (Watt)	Δ VO ₂ max (ml)	Reference
198	1.04	-	12	-26	_	-	_	37
27	1.00	_	12		-45	-	-	38
33	-	43	18	-34	-	-8	-278	39
30	_	39	24	+25	_	_	_	40
610	_	27	24	-	-	-	-	41
64	1.14	-	24	-44	-	-	-210	42

FEV₁, initial FEV₁ in litres or percentage predicted; follow up, duration of period over which exercise tolerance was assessed; Δ 6MWD, change in 6 minute walking distance in metres over the respective follow up period; Δ shuttle test, change in shuttle distance in metres over the respective follow up period; Δ Wmax and Δ VO₂max, change in maximal work load and oxygen consumption over the respective follow up period.

Decline in health status

It is well established that health status and health related quality of life—terms which will be used interchangeably here—do not correlate strongly with lung function parameters like FEV₁.⁴³⁻⁴⁵ This would suggest that the same physiological limitation may affect quality of life of COPD patients differently, and it justifies the use of quality of life outcomes in addition to clinical outcomes. Although the degree of association is not strong, there is a clear association between the two outcomes when a wider range of FEV₁ % predicted is considered.⁴⁶ Quality of life has also been shown to be related to exercise performance, ¹⁶ healthcare utilisation, ⁴⁷ ⁴⁸ and mortality.⁴⁹ ⁵⁰

Cross sectional studies have shown that quality of life deteriorates as lung function declines. In the early stages of COPD the decline in lung function might go unnoticed and might not affect quality of life. In a cross sectional study in COPD patients aged >64 years the progression of COPD from GOLD stages 0 to III did not correspond to a statistically significant difference in health status.46 However, an FEV1 of 50% predicted or less seemed to mark a threshold below which the St George Respiratory Questionnaire (SGRQ) total and impact score became clearly impaired.⁴⁶ This association between GOLD stage and SGRQ score was found after correction for differences in age, sex, living conditions, waste/ hip ratio, occiput wall distance, and number of co-morbidities, all variables that are related to quality of life. However, the variability in health status scores within a GOLD stage was very high.

The ISOLDE study was the first prospective study to show that COPD related quality of life deteriorated linearly over a 3 year period.¹⁷ As a result, reducing the decline in quality of life might become an informative end point in future long term trials. The ISOLDE study showed that the SGRQ total score worsened by 3.2 units per year in the placebo group compared with 2.0 units per year in the fluticasone propionate group—a statistically significant reduction in the rate of deterioration of 37%. However, in the BRONCUS trial, which included less severe COPD patients, the SGRQ total score worsened only by approximately 1.3-1.5 units a year in both the placebo and active treatment group during the last 2 years of the trial after an initial improvement in the first year.33 51 The difference between the two trials is perhaps related to the fact that lung function was optimised with a course of oral prednisolone before the patients were randomised into the ISOLDE trial, which was not the case before the start of the BRONCUS trial. Whether the decline in quality of life accelerates when lung function worsens is not yet clear. In the ISOLDE trial a weak but significant association between baseline FEV1 and deterioration in the quality of life was only found for the SGRQ impact score.17

It has recently been shown that a reduction in the decline in quality of life might be causally related to a reduction in the number of exacerbations.⁵² As COPD worsens the frequency and severity of exacerbations might increase, and

this increased frequency might be associated with an accelerated decline in quality of life.52 The centrality of exacerbations in the patients' lives is illustrated in a qualitative study in which patients told a number of near death stories that described severe exacerbations as life changing.⁵³ COPD patients with higher exacerbation rates are reported to have a poorer quality of life than those with lower exacerbation rates. 54 The impact of an exacerbation on quality of life may be considerably longer than just a temporary impairment of quality of life. In a study that evaluated the time course of recovery of health status after an exacerbation treated with antibiotics, it was shown that SGRQ scores continued to improve up to 26 weeks after the initial exacerbation that marked the end of the observation period. 17 The greatest improvement was seen during the first 4 weeks, but improvements during the second and third month were still above the 4 units considered to be clinically relevant. When patients had a further exacerbation within the 6 months after the initial exacerbation, health status did not improve much. These results are in line with those of Seemungal et al55 who reported that, 91 days after the onset of an exacerbation, peak expiratory flow rates and symptoms had not fully recovered in about 7% and 5% of exacerbations, respectively.

In 1 year or shorter term trials quality of life after treatment often improved, even in the placebo group. 7 8 This makes reduction in the declining quality of life a useless end point in these short term trials. A more relevant outcome in short term trials might be the proportion of patients reaching improvement beyond the minimal clinically important difference (MCID).56 The concept of MCID was introduced to improve the potential use of information provided by quality of life instruments.⁵⁶ This is important in order to distinguish statistically significant effects from clinically relevant effects, because small but clinically irrelevant effects can be statistically significant when the sample size is sufficiently large. Further work on the interpretability of scores and changes in scores is necessary. This includes further work on the association between the MCID and the severity of COPD. It also includes a stronger foundation for the interpretation of the MCID using external reference measures of clinically relevant change such as the occurrence of exacerbations or the use of healthcare services like hospital admissions. The proportion of patients improving by at least the MCID has become an increasingly used end point in addition to the mean change in scores, because it can easily be related to costs in order to report the cost effectiveness ratio ("costs per additional patient with improved quality of life"). 57-60 In addition to the proportion of improved patients, it is important to report the proportion showing deterioration greater than or equal to the MCID. Because of the natural decline in quality of life, the proportion of patients reaching at least the MCID is not a useful end point in long term studies.

Most long term clinical drug trials and many methodological studies on quality of life in COPD have used the SGRQ. In addition to this questionnaire, several other reliable and valid disease specific quality of life measures such as the Chronic Respiratory Disease Questionnaire (CRQ) and the Quality of Life for Respiratory Illness Questionnaire (QoLRIQ) have proved to be sensitive to changes in quality of life that result from treatments including pharmacotherapy. 61-63 The CRQ is a little different in that its original version includes an individualised dyspnoea domain. This means that, at baseline, patients identify five important daily activities and report their degree of dyspnoea on these same five activities at baseline and during follow up visits. This was done to improve the instrument's responsiveness. Since then, there has been an ongoing debate about the lack of standardisation, the limited comparability of CRQ scores, and the difficulty in eliciting activities that patients perform regularly to allow evaluation at repeated measurements. In a head-to-head comparison of the individualised and standardised approach it was recently shown that the individualised approach was more responsive to change, which has important consequences for reducing the required sample size.64 Moreover, while the SGRQ is the most widely used scale for health quality of life, it is still doubtful whether its complexity allows incorporation in daily practice. If decline in health status is to be assessed as an outcome measure in routine care for COPD, there is a need for a simple but reliable

Especially when designing long term studies with many repeated measurements of quality of life, it is important to note that several mechanisms that have been addressed by the term "response shift" might prevent instruments from detecting an objective deterioration or improvement in quality of life. Patients may recalibrate the measurement continuum or scale back their reference point against which their current health is compared. Hence, the quality of life scale no longer has the same psychological anchors as before. Patients may change their opinion about the relative importance of the aspects of quality of life or may redefine or reconceptualise quality of life. All these processes are phenomena of normal adaptive self-regulation that might seriously affect changes in self-reported quality of life scores over time.

Increase in healthcare utilisation and costs

In a review paper on the economic aspects of COPD the medical care costs per COPD patient in 2000 were reported to vary from \$US930 in Sweden to \$US2631 in the UK.66 It has been estimated that the annual costs per COPD patient are approximately three times the costs of asthma.^{67 68} Several cross sectional studies have shown a strong association between the increase in healthcare resource utilisation and the progression of COPD.69-72 Costs are disproportionately distributed. A relatively small proportion of severely ill patients with COPD are responsible for a substantial share of the total healthcare costs of COPD patients. $^{\rm 67\ 73\ 74}$ COPD patients with co-morbidities are particularly costly and the Confronting COPD Study has estimated that patients with co-morbidities account for 30-57% of the COPD population.⁷² Hospital admissions and medications are the two major cost drivers,66-74 but which of the two ranks first depends on the severity of COPD and the country or region under consideration.

Unscheduled care is the major portion of the direct medical costs. Exacerbations are a key driver of the costs of unscheduled care. In severe COPD, exacerbations are more frequent⁷⁵ and the severity of the exacerbation itself is also related to the underlying severity of the disease.^{76 77} Admission to hospital for a COPD exacerbation may become

necessary when initial outpatient treatment has failed. Failure rates of 12–21% have been reported.⁷⁵ ⁷⁸ The Spanish study by Miravittles *et al* ⁷⁸ estimated the costs associated with treatment failure to be eight times higher than the average cost of a successfully treated exacerbation. Estimates of the costs of exacerbations vary from \$US160 in a German study to \$US797 in a Dutch study (price level 2002)*.^{76–81} Costs rise with the severity of the exacerbation,⁷⁷ ⁸⁰ ⁸¹ with the Dutch study reporting that the mean costs of severe exacerbations were seven times the costs of moderate exacerbations and 47 times the costs of mild exacerbations.⁸¹

It is obvious that the potential for cost savings is largest for treatments that slow down disease progression, prevent exacerbations, and treat exacerbations effectively. From an economic perspective, it is most efficient to target the new and increasingly expensive treatments to the group of COPD patients who are at increased risk of being admitted to hospital for an exacerbation. It is therefore necessary to identify the risk factors of hospitalisation. A large number of potential risk factors for exacerbations and hospital admissions have been identified in the literature including:

- low body mass index and weight loss;
- number, type and severity of co-morbidities;
- history of frequent exacerbations;
- history of high healthcare resource use;
- number and type of medications;
- use of oxygen therapy;
- being a smoker and number of pack-years;
- housing status (living alone);
- lung function;
- degree of dyspnoea and other symptoms of COPD;
- impaired health status/quality of life;
- low physical activity;
- abnormal blood gases.

These predictors are not mutually exclusive but may be interrelated.

Table 4 gives an overview of the cost of improving one additional patient's quality of life as reported in some economic evaluations. Even though the studies use different quality of life instruments, a table like this might provide a general impression of the relative cost effectiveness of these treatments.

Prospective economic evaluations usually present the proportion of patients using a particular type of healthcare service and the mean healthcare utilisation and costs per patient per year. Statistical tests of the differences between treatments in the use and costs of separate types of healthcare services should be interpreted with care as substitution effects are common. A treatment that reduces the number of hospital days may increase the number of physician visits. Showing that an intervention reduces the use of healthcare services, especially hospital admissions, requires large sample sizes as the mean annual number of admissions is low in patients with stable COPD at study entry (between 0.11 and 0.24 per patient per year). 51 61

In addition to testing whether differences in costs are statistically significant, the aim of every economic evaluation is to obtain an unbiased estimate of the magnitude of this difference. This is particularly challenging because of the combined occurrence of a few characteristics that are typical

^{*}Costs were converted to 2002 \$US using harmonised indices of consumer prices published by EUROSTAT to correct for inflation and purchasing power parities as published by OECD to convert national currencies into \$US.

 Table 4
 Costs per additional patient with relevant improvement in quality of life

Interventions	Time scale	Year	Costs per additional patient with improved QoL ≥MCID	QoL measure
Tiotropium v ipratropium (MDI) ⁶⁰ Tiotropium v ipratropium (44% MDI/56% DPI) ⁶⁰	1 year 1 year	2001 2001	Euros 1084 Euros 289	SGRQ total SGRQ total
Salmeterol v placebo ⁵⁹ Early treatment with fluticasone v placebo ⁵⁸	4 months 1 year	Not reported 1998	UK £497 US \$1674	SGRQ total CRQ dyspnoea
Detection+early treatment with fluticasone v placebo ⁵⁸	1 year	1998	US \$4361	CRQ dyspnoea
Rehabilitation (2 months inpatient and 4 months outpatient) v conventional community care ⁵⁷	6 months	1989	Can \$28993 Can \$38270 Can \$47548 Can \$51027	CRQ mastery CRQ emotional CRQ dyspnoea CRQ fatigue

MCID, minimal clinically important difference; MDI, metered dose inhaler; DPI, dry powder inhaler; QoL, quality of life; SGRQ, St George's Respiratory Questionnaire; CRQ, Chronic Respiratory Questionnaire. 61-64

for cost data.82 Firstly, the variable of interest in cost analysis is the cumulative cost over a certain time period whereas clinical and quality of life data are often measured at several points in time. Secondly, costs tend to vary widely over time within one patient. A patient who is being hospitalised during time interval t will have very high costs over this interval while costs over t-1 and t+1 can be very low. Thirdly, resource use data are usually characterised by severe skewness of the underlying frequency distribution—that is, a few patients will have very high costs and the majority of patients will have limited costs. The most costly patients are usually the most severely ill patients who are most likely to drop out from a clinical trial, especially from a placebo control group. In statistical terminology, drop out is not "completely at random" but either "at random" or "informative". 19 82 It is important to stress that these types of drop out complicate both the clinical analyses and the economic analyses. Instead of ignoring the problem and not reporting the nature of the drop out or the methods that have been used to estimate the costs of drop outs, there are now socalled principled methods of imputation available that outperform the naïve methods such as complete case analysis, linear extrapolation, last value carried forward, and others. "Outperforming" means that they produce a less biased estimate of the mean costs and a more accurate estimate of the uncertainty around the mean.83 One of these principled methods—multiple imputation using propensity scores—has recently been applied in the cost effectiveness study of tiotropium versus ipratropium. 60 In a separate paper on the same study, the impact of different imputation methods on the cost estimates was shown.8-

In multinational economic evaluations the challenge of estimating cost effectiveness is even greater as it is far more difficult to transfer costs between countries than to transfer clinical data. The pure biological effect of a drug can be expected to be the same, irrespective of the country, but resource use and cost data are setting specific. They depend on local treatment patterns, which may be influenced by culture and traditions, the structure of the healthcare system, financial and other incentives. 85 Differences in absolute prices and, more importantly, relative prices may cause the cost effectiveness of a treatment to differ from country to country. As a consequence, simply pooling data from different countries without any adaptation of the data to the local setting or any testing of whether the effect of the treatment on resource utilisation is the same across countries is theoretically not sound. However, it is often unavoidable as the trials usually lack statistical power to do separate analyses per country and the adaptation of the resource use data to the local setting should be pre-specified in the statistical analysis plan of an economic evaluation.

An alternative or, even better, an addition to the empirical economic evaluations is modelling the cost effectiveness. Models can be populated with different country-specific treatment patterns and costs and they can be used to extrapolate trial results to a longer time horizon and include quality adjusted survival. In particular, stochastic models that address the uncertainty surrounding model based cost effectiveness estimates can be useful.⁸⁶

CONCLUSIONS

Recently conducted long term trials have contributed substantially to our insight into the progression of COPD. The main methods used are annual decline in FEV_1 , decline in functional and exercise capacity, decline in health status, and use of healthcare resources. Each of these variables is associated with significant problems that are largely related to selective drop out occurring in long term studies. This puts a particular burden on the statistical analysis and may require the use of more sophisticated models of analysis.

ACKNOWLEDGEMENTS

The authors acknowledge the expert secretarial assistance of Mrs C Ausseloos, Mrs E Lahousse, and Mrs A Vandeborne.

Authors' affiliations

M Decramer, R Gosselink, Respiratory Division, University Hospital, Katholieke Universiteit Leuven, 3000 Leuven, Belgium

M Rutten-Van Mölken, Institute for Medical Technology Assessment, Erasmus Medical Centre Rotterdam, The Netherlands

J Buffels, Department of General Practice, Katholieke Universiteit Leuven, 3000 Leuven, Belgium

O V Schayck, Department of General Practice, University of Maastricht, Maastricht, The Netherlands

P-A Gevenois, Department of Radiology, Erasme University Hospital, Brussels, Belgium

R Pellegrino, Centro di Fisiopatologia Respiratoria e dello Studio della Dispnea, Azienda Ospedaliera S Croce e Carle, 12100 Cuneo, Italy E Derom, Department of Respiratory Diseases, University of Ghent, Belaium

W D Backer, Department of Pulmonary Disease, University of Antwerp, Belaium

This workshop was supported by GlaxoSmithKline-Belgium.

REFERENCES

1 Pauwels RA, Lofdahl CG, Laitinen LA, et al. Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. N Engl J Med 1999;340:1948–53.

- 2 Vestbo J, Sørensen T, Lange P, et al. Long-term effect of inhaled budesonide in mild and moderate chronic obstructive pulmonary disease: a randomised controlled trial. Lancet 1999;353:1819-23.
- 3 Burge PS, Calverley PMA, Jones PW, on behalf of the ISOLDE study investigators, et al. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive oulmonary disease: the ISOLDE trial. BMJ 2000;320:1297-303.
- 4 The Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. N Engl J Med 2000;343:1902–9.
- 5 Vincken W, van Noord JA, Greefhorst APM, et al. Improved health outcomes in patients with COPD during 1 year's treatment with tiotropium. Eur Respir J 2002:19:209-16.
- 6 Casaburi R, Mahler DA, Jones PW, et al. A long-term evaluation of once-daily titropium in chronic obstructive pulmonary disease. Eur Respir J 2002;19:217-24
- 7 Szafranski W, Cukier A, Ramirez A, et al. Efficacy and safety of budesonide/ formoterol in the management of chronic obstructive pulmonary disease. Eur Respir J 2003;21:74-81.
- Calverley P, Pauwels R, Vestbo J, et al. Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. Lancet 2003;361:449-56.
- 9 Buffels J, De Gryse J, Heyrman J, et al. Office spirometry significantly improves early detection of COPD in general practice. The DIDASCO study. Chest 2004;**125**:1394–9.
- 10 Lundbäck B, Lindberg A, Lindström M, et al. Not 15% but 50% of smokers develop COPD? Report from the Obstructive Lung Disease in Northern Sweden studies. Respir Med 2002;97:115-22.
- Zielinski J, Bednarek M, et al. Early detection in a high-risk population using spirometric screening, Chest 2001;119:731-6.
- 12 Pauwels R, Buist A, Calverley P, on behalf of the GOLD Scientific Committee, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2000:**163**:1256-76.
- 13 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;158:1557-65
- 14 Pellegrino R, Rodarte JR, Brusasco V. Assessing the reversibility of airway obstruction. Chest 1998;114:1607-12.
- 15 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**:1042-50
- 16 Jones PW, Quirk F, Baveystock CM, et al. A self-complete measure of health status for chronic airflow limitation. The St George's Respiratory Questionnaire. Am Rev Respir Dis 1992;145:1321-7.
- Spencer S, Jones PW and the GLOBE Study Group. Time course of recovery of health status following an infective exacerbation of chronic bronchitis. Thorax 2003:**58**:589-93.
- 18 Dirksen A, Dijkman JH, Madsen F, et al. A randomized clinical trial of alpha1antitrypsin augmentation therapy. Am J Respir Crit Care Med 1999;160:1468–72.
- Little RJA, Rubin DA. Statistical analysis with missing data. New York: John Wiley and Sons, 1987.
- Siafakas NMP, Vermeire NB, Pride P, et al. Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. Eur Respir J 1995;8:1398–420.
- 21 Quanjer PH, Tammeling GJ, Cotes JE, et al. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. Eur Respir J 1993;6:5-40.
- 22 Rennard S, Decramer M, Calverley PMA, et al. Impact of COPD in North America and Europe in 2000: subjects' perspective of the Confronting COPD International Survey. Eur Respir J 2002;20:799-805.
- 23 Eaton T, Withy S, Garrett JE, et al. Spirometry in primary care practice: the importance of quality assurance and the impact of spirometry workshops. Chest 1999;116:416-23.
- 24 Mulder HH, vander Molen T, Postma DS, et al. Measuring FEV₁ in general practice. Eur Respir J 1996;9:121S.
- 25 Ponsioen BP, Hop WCJ, Vermue NA, et al. Characteristics of adults with persistent cough in general practice. Primary Care Respir J 2000;11:68–9.
 26 Shermer TR, Jacobs JE, Chavannes NH, et al. Validity of spirometric testing in
- a general practice population of patients with chronic obstructive pulmonary disease (COPD). *Thorax* 2003;**58**:861–6.
- 27 Jackson H, Hubbard R. Detecting chronic obstructive pulmonary disease using eak flow rate: cross sectional survey. BMJ 2003;327:653.
- 28 Macklem PT, Thurlbeck WM, Fraser RG. Chronic obstructive disease of small airways. Ann Intern Med 1971:167-77
- de Miguel Diez J, Izquierdo Alonso JL, Molina Paris J, et al. Reliability of chronic obstructive pulmonary disease diagnosis by primary care physicians and pneumologists in Spain. Predictive factors. Arch Bronconeumol 2003:**39**:203-8
- Burrows B, Earle RH. Course and prognosis of chronic obstructive lung disease. N Engl J Med 1969;280:397–404.
- $\textbf{Anthonisen N\~R}, \textbf{Connett JE}, \textbf{Kiley JP}, \textit{et al.} \textbf{ Effects of smoking intervention and}$ the use of anticholinergic bronchodilator on the rate of decline of FEV₁. The Lung Health Study. *JAMA* 1994;**272**:1497–505.
- Dirksen A, Holstein-Rathlou NH, Madsen F, et al. Long-range correlations of serial FEV₁ measurements in emphysematous patients and normal subjects. J Appl Physiol 1998;**85**:259–65.

- 33 Decramer M, Dekhuijzen PNR, Troosters T, et al. and the BRONCUS trial committee. The bronchitis randomised on NAC cost-utility study (BRONCUS): hypothesis and design, Eur Respir J 2001;17:329–36.
- Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. Appl Statist 1994;**43**:49–93.
- 35 Andreassen H, Vestbo J. Chronic obstructive pulmonary disease as a systemic disease: an epidemiological perspective. Eur Respir J 2003;46(Suppl.):2s-4s.
- 36 Marquis K, Debigare R, Lacasse Y, et al. Mid thigh muscle cross-sectional area is a better predictor of mortality than body mass index in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2002:166:809-13
- 37 Pinto-Plata VM, Cote C, Cabral H, et al. The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. Eur Respir J 2004:23:28-33
- Bestall JC, Paul EA, Garrod R, et al. Longitudinal trends in exercise capacity and health status after pulmonary rehabilitation in patients with COPD. Respir Med 2003;97:173-80.
- **Troosters T**, Gosselink R, Decramer M. Long-term effects of outpatient rehabilitation of COPD. A randomized trial. *Am J Med* 2000;**109**:207–12.
- Güell R, Casan P, Belda J, et al. Long-term effects of outpatient rehabilitation of COPD: a randomized trial. Chest 2000;117:976–83.
- National Emphysema Treatment Trials Research Group. A randomized trial comparing lung volume reduction surgery with medical therapy for severe emphysema. N Engl J Med 2003;348:2059-73.
- 42 Ries AL, Kaplan RM, Myers R, et al. Maintenance after pulmonary rehabilitation in chronic lung disease: a randomized trial. Am J Respir Crit Care Med 2003;**167**:880–8.
- Ferrer M, Alonso J, Prieto L, et al. Validity and reliability of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. Eur Respir J 1996;9:1160-6.
- Hajiro T, Nishimura K, Tsukino M, et al. Comparison of discriminative properties among disease-specific questionnaires for measuring health-related quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1998;158:1185–9.
- Rutten-van Mölken MPMH, Roos B, Van Noord JA. An empirical comparison of the St Georges Respiratory Questionnaire and the Chronic Respiratory Questionnaire in a clinical trial setting. *Thorax* 1999;54:995–1003.

 46 Antonelli-Inc C, Imperiale V, Bellia F, et al. Do GOLD stages of COPD severity
- really correspond to differences in health status? Eur Respir J 2003;22:444-9.
- Van den Boom G, Rutten-van Mölken MPMH, Tirimanna PRS, et al. Association between health-related quality of life and consultation for respiratory symptoms: results from the DIMCA programme. Eur Respir J 1998;**11**:67–72
- 48 Alemayehu B, Aubert RE, Feifer RA, et al. Comparative analysis of two quality-of-life instruments for patients with chronic obstructive pumonary disease. Value Health 2002;**5**:436–41.
- Fan V, Curtis JR, MacDowell M, et al. Using quality of life to predict hospitalizations and death in chronic obstructive lung disease. Chest 2002:122:429-36
- Domingo-Salvany A, Lamarca R, Ferrer M, et al. Health-related quality of life and mortality in male patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2002;**166**:680–5.
- Decramer M, Rutten-van Mölken MPMH, Dekhuijzen PNR, et al. Effects of Nacetyl-cysteine on outcomes in COPD. The Bronchitis Randomized on NAC Cost-Utility Study (BRONCUS). Lancet 2005 (in press).
- Spencer S, Calverley PMA, Burge PS, et al. Impact of preventing exacerbations on deterioration of health status in COPD. Eur Respir J 2004;23:698-702
- 53 Bailey Ph. Death stories: acute exacerbations of chronic obstructive pulmonary disease. Qual Health Res 2001;11:322-38.
- Seemungal TAR, Donaldson GC, Paul EA, et al. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1998;157:1418–22.

 Seemungal TAR, Donaldson GC, Bhowmik A, et al. Time course and recovery
- of exacerbations in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2000;**161**:1608–13.
- Hays RD, Wolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? *PharmacoEconomics* 2000.18.1215-9
- Goldstein RS, Gort EH, Guyatt GH, et al. Economic analysis of respiratory rehabilitation. Chest 1997;112:370–9.
- Van den Boom G, Rutten-van Mölken MP, Molema J, et al. The cost effectiveness of early treatment with fluticasone propionate 250 micrograms twice a day in subjects with obstructive airway disease: results of the DIMCA program. Am J Respir Crit Care Med 2001;164:1057-66.
- Jones PW, Wilson K, Sondhi S. Cost-effectiveness of salmeterol in patients with chronic obstructive pulmonary disease: an economic evaluation. Respir Med 2003;97:20-6.
- Oostenbrink JB, Rutten-van Mölken MPMH, Al MJ, et al. One-year costeffectiveness of tiotropium versus ipratropium to treat chronic obstructive pulmonary disease. *Eur Respir J* 2004;**23**:241–9.
- Jones OW, Bosh TK. Quality of life changes in COPD patients treated with salmeterol. Am J Respir Crit Care Med 1997;155:1283–9.
- Mahon JL, Laupacis A, Donner A, et al. Randomised study of n of 1 trials
- versus standard practice. *BMJ* 1996;**312**:1069–74. **Van Stel HF**, Maillee AR, Colland VT, *et al.* Interpretation of change and longitudinal validity of the quality of life for respiratory illness questionnaire (QoLRIQ) in inpatient pulmonary rehabilitation. *Qual Life Res* 2003;**12**:133–45.

- 64 Schunemann HJ, Griffith L, Jaeschke R, et al. A comparison of the original Chronic Respiratory Questionnaire with a standardized version. Chest 2003; **124**: 1421-9.
- Carver CS, Scheier MF. Scaling back goals and recalibration of the affect system are processes in normal adaptive self-regulation: understanding 'response shift' phenomena. Soc Sci Med 2000;50:1715–22.

 Ruchlin HS, Dasbach EJ. An economic overview of chronic obstructive
- pulmonary disease. PharmacoEconomics 2001;19:623–42
- Sullivan SD, Ramsey SD, Lee TA. The economic burden of COPD. Chest 2000;**117**:5–9S.
- Rutten-van Mölken MPMH, Postma MJ, Jore MA, et al. Current and future health care impact of asthma and COPD in The Netherlands. Respir Med 1999;93:779-87.
- Hilleman DE, Dewan N, Malesker M, et al. Pharmacoeconomic evaluation of
- chronic obstructive pulmonary disease. *Chest* 2000;118:1278–85.

 Jansson SA, Andersson F, Borg S, *et al.* Costs of COPD in Sweden according to disease severity. *Chest* 2002;122:1994–2002.
- Miravitlles M, Murio C, Guerrero T, on behalf of the DAFNE Study Group, et al. Costs of chronic bronchitis and COPD. A 1-year follow-up study. Chest 2003;123:784–91.
- 72 Wouters EFM. Economic analysis of the Confronting COPD survey: an overview of results. Respir Med 2003;97(Suppl C):S3-14.
 73 Grasso ME, Weller WE, Shaffer TJ, et al. Capitation, managed care, and
- chronic obstructive pulmonary disease. Am J Respir Crit Care Med 1998;158:133-8.
- Strassels SA, Smith DH, Sullivan SD, et al. The costs of treating COPD in the United States. Chest 2001;119:344-52.
- 75 Dewan NA, Fafique S, Kanwar B, et al. Acute exacerbations of COPD: factors associated with poor treatment outcome. Chest 2000;117:662-71.

- 76 Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. Chest 2000;117:398-401S.
- Andersson F, Borg S, Jansson SA, et al. The costs of exacerbations in chronic
- obstructive pulmonary disease (COPD). Respir Med 2002;96:700-8.

 78 Miravitles M, Murio C, Guerro T, et al. Pharmacoeconomic evaluation of acute exacerbatios of chronic bronchitis and COPD. Chest 2002:121:1449-55.
- Pechevis M, Fagnani F, Brin S, et al. Infections respiratoires récidivantes du sujet atteint de bronchite chronique obstructive: prise en charge médicale et coûts. Rev Mal Resp 1996;13:507–12.

 Rychlik R, Pfeil T, Daniel D, et al. Socioeconomic relevance of acute
- exacerbations of chronic bronchitis in the Federal Republic of Germany. A prospective cost of illness study. Dtsch Med Wochenschr 2001:126:353-9.
- Oostenbrink JB, Rutten-van Mölken MPMH. Resource use and risk factors of high cost exacerbations. Respir Med 2004;98:883-91
- Rubin D, Schenker N. Multiple imputation in healthcare databases: an
- overview and some applications. Stat Med 1991;10:595–8.

 Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. Health Economics 2005;(in press).
- Oostenbrink JB, Al MJ, Rutten-van Mölken MPMH. Methods to analyse cost data of patients who withdraw in a clinical trial setting. PharmacoEconomics 2003;**21**:1103–12.
- 85 Drummond M, Pang F. Transferability of economic evaluation results. In: Drummond M, McGuire A, eds. Économic evaluation in health care. Merging theory with practice. Oxford: Oxford University Press, 2001:256-76
- Briggs AH. Handling uncertainty in cost-effectiveness models PharmacoEconomics 2000;17:479-500.

Register now!

10th European Forum on Quality Improvement in Health Care 13-15 April 2005, ExCel Conference Centre, London For further information on how to register please go to: http://www.quality.bmjpg.com