REVIEW

Outcomes and markers in the assessment of chronic obstructive pulmonary disease

P.W. Jones* and A.G.N. Agusti[#]

ABSTRACT: **The clinical presentation of chronic obstructive pulmonary disease (COPD) is highly variable, reflecting the interaction of a complex range of pathological changes including both pulmonary and systemic effects.**

The consequences of COPD experienced by the patient (*i.e.* its outcomes) include: symptoms, weight loss, exercise intolerance, exacerbations, health-related quality of life, health resource use and death. No single measure can reflect the variety of pathological effects or adequately describe the nature or severity of COPD.

Currently, there are few validated markers for assessing COPD and evaluating the effectiveness of treatment. The forced expiratory volume in one second has been used as a global marker of COPD, but it does not fully reflect the burden of COPD on patients. New markers are needed to better characterise the full clinical spectrum of the disease and to guide the development and assessment of new and more effective therapies.

This article considers the distinction between outcomes and markers, the various ways in which markers are used and the need for new markers in the management of chronic obstructive pulmonary disease. The process of marker selection and validation is reviewed and potential new biological, physiological and symptomatic markers for chronic obstructive pulmonary disease are assessed.

AFFILIATIONS
*Respiratory Medicine, St. George's, University of London, London, UK.
#Servei de Pneumologia, Hospital Universitari Son Dureta, Fundació Caubet-Cimera, Palma de Mallorca, Spain.

CORRESPONDENCE
P.W. Jones
Respiratory Medicine
St. George's
University of London
Cranmer Terrace
London
SW17 0RE
UK
Fax: 44 208725955
E-mail: pjones@sgul.ac.uk

Chronic obstructive pulmonary disease (COPD) is defined currently by the American Thoracic Society (ATS) and the European Respiratory Society (ERS) as a disease characterised by airflow limitation, which is not fully reversible and produces systemic consequences. It is a multicomponent disease characterised by a range of pathological changes, which include mucus hypersecretion, airway narrowing and loss of alveoli within the lungs, and loss of lean body mass and cardiovascular effects outside the lungs. There is extensive heterogeneity among patients with COPD in terms of their clinical presentation, disease severity and rate of disease progression. Two general terms are used in the context of measurement of the effect of disease: outcomes and markers. It is useful to retain the term "outcome" for the consequences of the disease as experienced by the patient. In COPD, these would include symptoms, weight loss, exercise intolerance, exacerbations, impaired health-related quality of life, increased health resource use and death. A marker is a measurement known to be associated with a clinical outcome. Thus, exercise capacity, as tested in a laboratory, is a marker of the patient's exercise intolerance in daily life, and health status scores provide a marker of the patient's health-related quality of life.

In the absence of other widely accepted and validated markers, lung function measurement and, specifically, the forced expiratory volume in one second (FEV_1), has been used as a global marker for all the pathophysiological changes in COPD. However, measures of lung function correlate only poorly with the severity of dyspnoea and other symptoms [1], and COPD has a range of extrapulmonary effects [2]. For that reason, lung function measurements by themselves are not adequate to describe the impact of COPD or assess the effectiveness of therapeutic interventions.

New markers are needed to allow a more complete and clinically relevant assessment of

COPD, reflecting the substantial variation in the way in which the disorder presents in different patients. They may enable better phenotyping of different types or patterns of COPD and help improve assessment of disease severity, response to therapy and monitoring of disease progression. They may also assist in new drug development processes. Conversely, the lack of appropriate markers and outcomes, or use of inappropriate ones, can present an obstacle to the development of new and effective therapies [3]. The process of identifying appropriate markers and outcomes is not straightforward and needs to reflect the needs of various parties with differing priorities (table 1). Validation of these measures is time-consuming and involves a substantial commitment of resources. However, this is critical to gaining a better understanding of the pathophysiology of COPD and the development of reliable, comprehensive and evidence-based assessments of the effectiveness of therapeutic interventions.

DEFINING OUTCOMES AND MARKERS

The meaning of the term outcome can vary according to the setting in which it is used. This article is concerned with clinical outcomes that are consequences of the underlying disorders in COPD and that are experienced directly by patients. These include disability, poor health, reduced quality of life and death. Outcomes can take on other meanings in

TABLE 1	Requirements of markers in chronic obstructive pulmonary disease for different groups, some of which, but not all, are shared
Group	**Key requirements**
Family physicians	Simplicity
	Specificity
	Reliability
	Clarity
Hospital specialists	Sensitivity
	Specificity
	Reliability
	Clinical utility
Payers	Cost-effective
	Clarity
Pharmaceutical industry	Practicality
	Sensitivity
	Generalisability
	Reliability
	Simplicity
Regulatory authorities	Sensitivity
	Specificity
	Reliability
	Clarity
	Clinical utility
	Generalisability
Scientists	Reliability
	Sensitivity
	Specificity
	Biological credibility
Patients	Clarity
	Simplicity

medicine, and these will be discussed below. For many disorders there is often a lack of clinical outcomes that can be measured easily within routine practice or in the setting of a clinical trial. In chronic conditions, such as COPD, the effect of the underlying pathological processes takes many years to develop. As a consequence, studies designed to evaluate the effects of treatment directly on clinical outcomes may need to run for a long time, which might delay the development of new treatments.

In conditions where direct measurement of treatment effects on clinical outcomes is not feasible, markers are needed. A marker has been defined as a clinical measurement that is associated with, and believed to be related pathophysiologically to, a clinical outcome [4, 5]. The term marker is used in a number of ways, so it is important that there is clarity in its use, since its meaning may change with its context.

Diagnostic marker
In this context, a marker is used as a dichotomous variable (*i.e.* it is either present or absent). The marker is usually measured on a continuous scale, but a threshold value is used to define presence or absence of the clinical state; for example, α_1-antitrypsin level or FEV1 when used for COPD diagnosis.

Measure of disease severity
In this role, the marker may describe more than two levels of disease severity, or stage, into which measurements are categorised according to predefined ranges. The chosen ranges for these categories may or may not be evidence based. Examples include: body mass index (BMI) or FEV1, as used in ERS, ATS guidelines and Global Initiative for Chronic Obstructive Lung Disease (GOLD) staging.

Marker of disease progression
Here, the measurement may be used as a continuous variable, without categorisation, *e.g.* rate of decline of FEV1 or rate of decline of health status score. Other markers may be categorised, *e.g.* presence or absence of hypoxia, presence of pulmonary hypertension or loss of lean body mass.

Marker of treatment effect
These are the familiar markers used to measure response to treatment (*e.g.* dyspnoea score, lean body mass, exercise capacity, health status, FEV1, *etc*.). Confusion around terminology can arise in this setting, because in clinical trials these markers are often termed ''outcome variables'', ''clinical end-points'' and ''clinical trial end-points''. These are, of course, outcomes from the trial and not clinical outcomes.

Biological markers and biomarkers
A widely used definition endorsed by the National Institutes of Health in the USA is that a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to therapeutic interventions. In this context, biomarkers are generally considered to be substances rather than physiological measurements, such as FEV1 or exercise capacity. An example in COPD is exhaled nitric oxide as a biomarker for an inflammatory process in the airways.

Surrogate marker

It will have become clear that all markers are surrogate markers for the outcome of interest, but this term applies when a marker is used as a substitute for the marker of primary interest, *e.g.* diffusing capacity as a surrogate marker for the presence of emphysema. In other words, a surrogate marker is one marker being used in place of another marker.

Markers used for categorising severity

It should be noted that, in a number of the applications described above, threshold values are used to partition measurements of a marker made on a continuous scale into defined categories, for example, mild/moderate/severe, *etc.* These thresholds should be set using validated methodologies, but initially they have often been set rather arbitrarily, with validation of the categories being carried out later. Categorisation of COPD severity using the FEV_1 is a clear example of the latter approach. Initially the chosen categories were rather arbitrary, but they have proved to have practical value for research and clinical practice, which is a component of validation. This example also illustrates the need to use markers of severity carefully. Staging criteria based upon the FEV_1 categorise COPD patients only in terms of the degree of airflow limitation. They do not categorise COPD severity, since that is multicomponent in nature.

RELATIONSHIP BETWEEN AN OUTCOME AND ITS MARKERS

An outcome may have multiple markers; for example, BMI, FEV_1 and exercise capacity are all independent predictors of mortality. Furthermore, the relationship between an outcome and its marker may be modified by factors internal to the patient, such as the presence of comorbid conditions or external factors, such as the level of family or social support and access to healthcare.

Within medicine in general, there are few markers that are so well characterised and understood that they can substitute effectively for a clinical outcome, and be used as a surrogate for the outcome. One of the rare examples is hypertension, where there is a wealth of evidence demonstrating that elevated blood pressure is a cause of cardiovascular and cerebrovascular disease, and that treatment to lower elevated blood pressure reduces morbidity and mortality from outcomes such as myocardial infarction and stroke [6]. Consequently, a reduction in blood pressure has become accepted as a treatment end-point in the prevention of these clinical outcomes. In other words, blood pressure reduction becomes both the therapeutic target and the proximate outcome for the patient, since it lowers the risk of developing other outcomes. In the management of diabetes, glycosylated haemoglobin takes on a similar role. Even in such cases, it is important to remember that the correlation between marker and outcome is not perfect. In fact it is probabilistic and may be situation-specific. For instance, antihypertensive therapy with angiotensin-converting enzyme inhibitors has greater benefits in reducing cardiovascular mortality and morbidity compared with diuretics, despite comparable reductions in blood pressure.

The strength of evidence linking a clinical outcome and its marker evolves over time as part of an ongoing process of validation. In many cases, the correlation between marker and outcome is not well defined and it is important to realise that the extent of validation of a marker is inversely related to its degree of novelty. A novel marker may reflect an important new insight into underlying disease mechanisms, but may not yet have been shown to correlate closely with a relevant clinical outcome. For example, inflammation is increasingly recognised as an important contributory factor in coronary artery disease and its associated cardiovascular outcomes. There are a variety of factors such as C-reactive protein, interleukin (IL)-6 and tumour necrosis factor-α (TNF-α) that are known to be directly involved in vascular inflammation, and elevated levels of these proteins correlate with increased risk of cardiovascular disease [7, 8]. However, unlike blood pressure and stroke, there is no direct evidence to demonstrate that treatment-associated reductions in these factors correlate with improvements in clinical outcomes. Whilst markers of inflammation cannot, at present, substitute for clinical outcomes, they can still help by advancing the understanding of coronary artery disease and assessing potential benefits of anti-inflammatory agents in the management of cardiovascular disease. Another problem is that markers may be very specific and apply to a particular part of a disease pathway. Many chronic diseases are complex, with multiple pathways, so a biological marker may reflect activity in only one of a number of different and important pathways.

THE NEED FOR NEW MARKERS IN COPD

The pathogenesis of coronary artery disease has been investigated extensively and is relatively well defined. COPD, by comparison, has been studied much less and the underlying disease mechanisms are less well understood. COPD is characterised by a range of pathological changes, including airflow limitation and inflammation, mucociliary dysfunction, structural changes in the airways and extrapulmonary systemic effects [2, 9–12]. Reflecting the multicomponent nature of the disorder, there is extensive heterogeneity among patients with COPD in terms of clinical presentation, disease severity and rate of disease progression. It is increasingly apparent that a single marker is unlikely to be predictive of clinical outcome in all patients with COPD, given the diverse range of pathological mechanisms involved. Furthermore, with the variable clinical presentation of COPD, a single outcome is unlikely to provide a full assessment of the impact of COPD across all patients. Despite this limitation, definitions of COPD from organisations including the ERS, ATS and GOLD [13–15] have focused almost exclusively on the lungs, although the ERS/ATS definition has recently been modified to include reference to the significant systemic consequences of COPD [16]. Furthermore, staging and prognosis of COPD is currently determined solely on the basis of lung function measurements, principally FEV_1. In effect, FEV_1 has come to be used as a global marker for all the pathophysiological changes in COPD, even though a number of the changes are extrapulmonary.

The role of FEV_1 as a global marker of COPD is reinforced by the requirements of regulatory authorities. For instance, the Committee for Proprietary Medicinal Products of the European Agency for the Evaluation of Medicinal Products states that the primary end-point in clinical trials should reflect the clinical

benefit and should include FEV1 as a measure of lung function as well as a measure of symptomatic benefit [17]. The ability of therapies to prevent the progression of COPD is thus judged in terms of effects on a single marker that is relevant to only some of the pathophysiological processes in COPD and so is unlikely to provide an accurate assessment of the overall clinical effect.

There is another risk in placing reliance solely upon the FEV1 as a marker of COPD. Bronchodilators are central to current strategies for managing COPD, and FEV1 is a reliable marker of a principal clinical effect of these therapies. It is possible that new agents that act by mechanisms other than bronchodilation might have little or no effect on FEV1, but significantly improve clinical outcomes such as mortality and hospitalisation. For example, pulmonary rehabilitation has been shown to improve exercise tolerance and symptoms of COPD including breathlessness and muscle fatigue, without modifying FEV1 or other lung function parameters [18–20]. Treatments that might have the potential to substantially improve the management of COPD might be denied to physicians and patients, if based solely on the current regulatory requirement that treatments should improve FEV1 [17].

The limitations of relying on a single marker to describe the progression of a multicomponent disease are compounded by a number of other considerations that, taken together, highlight the need for new and additional markers in COPD. For example, the use of FEV1 to assess treatment efficacy seems paradoxical, since COPD is diagnosed on the basis of low FEV1 that is poorly responsive to bronchodilator therapy. In other words, new treatments for COPD are required to modify a marker in a population that has been selected on the basis of the unresponsiveness of that same marker to established COPD therapies. Given this paradox, it is perhaps unsurprising that many trials of bronchodilators and anti-inflammatory agents have shown only marginal efficacy in terms of improvements in FEV1 [21–27].

It is important to remember that FEV1 is a marker, and not a clinical outcome, of COPD. Whilst the FEV1 has been shown to correlate with mortality and health status, these correlations are weak, and it is not until FEV1 falls to <50% of the predicted level that mortality begins to rise substantially. Furthermore, in patients where the percentage of predicted FEV1 has fallen to very low levels, this measure has little predictive value [14, 28]. Evidence suggests that other measures may be better indicators of disease progression than FEV1. One study, for instance, found that staging of COPD based on the severity of dyspnoea was more predictive of 5-yr survival than categorisation according to percentage of predicted FEV1 [1]. Body weight [29, 30], exercise capacity and health status have also been shown to correlate with mortality in COPD [31–34].

The concept of a single global marker has the attraction of simplicity and convenience, but may not be appropriate to a complex, multicomponent disorder, such as COPD. Additional markers and outcomes are needed to provide a more comprehensive and clinically meaningful assessment and so provide a more informed basis for treatment decisions. In particular, markers related to inflammatory processes, structural changes and systemic effects could yield valuable information to complement that provided by FEV1 for airflow limitation. Given that COPD is a progressive disorder, it may be that certain markers and outcomes are more relevant and useful at particular disease stages.

SELECTION AND VALIDATION OF NEW MARKERS IN COPD

It is much easier to criticise the inadequacies of current practice than to identify and take the steps needed to overcome those shortcomings. The complexity of COPD, in terms of both its pathogenesis and clinical presentation, means that there is a wide range of potential markers that might provide valuable information. Selecting which of these might prove both practical and effective requires balancing a broad range of factors and the differing needs of various groups involved in the investigation and management of COPD (table 1). For example, key requirements of research scientists and hospital specialists will be that the marker has the potential to extend their knowledge and understanding of COPD. Ease of measurement or interpretation of the marker will generally be of secondary importance for these groups. In contrast, the simplicity of the measure is a primary consideration for family physicians. Ideally, a marker used routinely in the primary care setting should require no special resources or processes beyond those typical of a patient–doctor consultation. Interpretation of the results should be clear and not require specialist knowledge, and patients should be able to understand their implications. From the perspective of the pharmaceutical industry, markers need to have sufficient practicality to allow consistent application in clinical trials and the clarity to determine and illustrate potential benefits in routine practice of the agent under investigation. Reliability (*i.e.* the ability to perform the same way in different settings) and generalisability are key requirements for regulatory authorities, so that trials with different therapies can be compared on an objective basis.

A number of key attributes are required of a new marker. 1) Relevant. The marker should relate to underlying disease mechanisms and/or the well-being of the patient. 2) Sensitive. The marker should be able to detect clinically important differences and be predictive of morbidity and mortality. In this respect it should perform the same way across all patients (discriminatory ability), as it does within patients (evaluative ability). 3) Selective and specific. The marker should not be influenced by confounding effects such as comorbidities or other factors. 4) Reliable. The marker should perform consistently in different settings, when administered to different patients by different healthcare professionals or research workers. 5) Internal consistency. Questionnaires and similar instruments, such as diary cards, should have good internal consistency. 6) Repeatable. Measurements should be stable and only change if the underlying disease changes. 7) Interpretable. The measurements obtained with the marker (for example, questionnaire scores) should translate into clinically interpretable results. 8) Simple. Testing and analysis should be accommodated easily within existing routine practice. For biomarkers, assays should not require procedures or equipment beyond those that are standard in a clinical laboratory. Noninvasive techniques are generally more convenient and acceptable for patients. 9) Cost-effective. The cost of monitoring should be more than offset by the savings resulting from improved management.

This is a very wide range of requirements and is difficult for any single marker to meet, so a degree of compromise is inevitable. In practice, the extent to which a marker conforms to this profile of attributes evolves over time as part of its validation [35]. With a new marker it is unlikely that reproducibility and predictive power will be clearly established, but its relevance to an underlying disease mechanism may promise an improved understanding and more comprehensive assessment of the disorder that can be established over time. As noted above, the degree of novelty in a marker is inversely related to the extent to which it is validated [4].

The validation processes of a marker for use in research and clinical trials can be considered to fall broadly into three phases (table 2). The first phase involves demonstrating that it is relevant to the natural history of the disorder and can plausibly be expected to predict clinical outcomes. Changes in the frequency and/or magnitude of the marker should correlate closely with clinical outcomes. In the second phase of validation, treatments known to be effective in the treatment of the particular condition should be shown to have effects on the marker consistent with an improvement in clinical outcomes. In studies comparing different treatments, the size of effect should correlate with changes in other related clinical outcomes or validated markers. The final phase of validation requires a demonstration that treatment-related changes in the marker correlate with positive changes in clinical outcomes. The validity of the marker should be assessed in treatments that act *via* different mechanisms. In COPD, for instance, the validity of the marker should also be assessed for both bronchodilator and anti-inflammatory agents. This phase should also assess whether the validity of the marker is restricted to a particular disease stage or covers the complete disease life cycle [35].

POTENTIAL NEW MARKERS FOR THE ASSESSMENT OF COPD

Whilst there are currently very few well-validated markers, there are a large number of candidate markers that could potentially be valuable for the assessment of COPD. A number of these are described in table 3 and are discussed below, but this list should not be considered exhaustive. For clarity, we have grouped them into different subgroups.

Biological markers

Validation of new biological markers is a difficult and time-consuming process. Given that airway inflammation is a central component in the pathogenesis of COPD. Inflammatory cells and mediators would present a logical target as potential markers for disease monitoring and the assessment of therapeutic interventions. COPD is characterised by neutrophilia [36] and increased levels of inflammatory mediators, including IL-6 [37], IL-8 [38, 39], TNF-α [37, 39] and leukotriene B$_4$ [40, 41]. Airway inflammation occurs in COPD, but also in smokers without COPD [42, 43] and in patients with COPD who do not currently smoke [44, 45]. For this reason, the presence of airway inflammation may not be sufficiently specific to act as a marker of COPD.

Studies have tried to establish if COPD is associated with a particular pattern of airway inflammation distinct from that seen in smokers without COPD or other conditions including asthma [46]. In a comparison of inflammatory cells in the peripheral airways of smokers with and without COPD, higher levels of CD8+ T-lymphocytes were found in biopsies from those patients with COPD. However, the levels of neutrophils, macrophages and CD4+ T-lymphocytes were similar in patients with or without COPD [47]. There are also differences in the pattern of inflammation in COPD as compared with asthma. In a comparison of patients with a history of either COPD or asthma and with similar degrees of fixed airflow obstruction and airway hyperresponsiveness, the patients with a history of COPD had significantly more neutrophils and significantly fewer eosinophils in sputum and bronchoalveolar lavage fluid, as well as a lower ratio of CD4+/CD8+ T-lymphocytes [48]. Thus, while it appears that COPD may be associated with a particular profile of inflammatory changes, a better understanding of the inflammatory mechanisms in COPD is needed to determine if any inflammatory cell or mediator, either alone or in combination, is sufficiently specific to act as a disease marker for COPD.

Whilst much attention has been paid to biomarkers from the lungs, biomarkers need not necessarily be present in the

TABLE 2	Phases of marker validation process for therapeutic trials (modelled on [35])	
Classification	**Description**	**Validation**
Phase 1	Natural history marker	Demonstration of relationship between frequency/magnitude
	Reflect underlying disease mechanisms	of marker and clinical outcomes
	Predict clinical outcomes independent of treatment	
Phase 2	Biological activity marker	Phase I/II trials demonstrating appropriate effects on the
	Responds to therapy	marker with therapeutic intervention
	Change in frequency/magnitude of marker corresponds	Dose-dependent effects of treatment on the marker
	with therapeutic potency	
Phase 3	Marker of therapeutic efficacy	Demonstration that the marker predicts clinical outcomes and
	An early change in the marker predicts clinical outcome	responds to therapy in outcome trials
	in the context of active therapy	Validation should establish whether the marker is applicable to
	Ideally, marker could be used as a surrogate endpoint	all disease stages and all interventions
	for clinical outcomes	

TABLE 3 Potential markers for the assessment of chronic obstructive pulmonary disease

Biological markers	Physiological markers	Symptomatic markers
Expectorated cellular markers of inflammation	**Markers of lung function**	**Symptoms**
Neutrophils, macrophages, eosinophils, mast cells, lymphocytes	Inspiratory capacity, $D_{L,CO}$	MRC Respiratory Questionnaire
Expectorated soluble markers of inflammation	**Physiological tests of small airway obstruction**	**Breathlessness**
TNF-α, IL-8, ECP, MPO	Lung hyperinflation	MRC Dyspnoea scale, Borg scale, BDI/TDI, UCSD dyspnoea scale
Expired gases	**Rate of decline of lung function**	
NO, CO, H_2O_2	**Exercise testing**	**Disease-specific health status (health-related quality of life)**
Expired air condensate	6-min walk test	CRDQ, SGRQ, BPQ, PFSDQ, PFSS, CCQ
LTB_4, cytokines, aldehydes	**Bronchial hyperreactivity**	
Peripheral blood markers	**Skeletal muscle function**	**Generic health status**
Activated neutrophils, TNF-α, soluble TNF receptors, IL-6, IL-8, CRP	**Lean body mass**	SF-36, NHP, EQ-5D
	Imaging	**Cognitive function**
Sputum protease/anti-protease levels	CT scan, PET, hyperpolarised gas MR	
HNE, MMPs, α_1-AT, SLPI, TIMPs	**Exacerbations**	
EGF levels in sputum	Rate	
Urine markers	Type	
Markers of matrix degradation, *e.g.* desmosine	**Quantification of luminal airway mucus**	
	Measures of pulmonary hypertension	

TNF-α: tumour necrosis factor-α; IL: interleukin; ECP: eosinophilic cationic protein; MPO: myeloperoxidase; NO: nitric oxide; CO: carbon monoxide; H_2O_2: hydrogen peroxide; LTB_4: leukotriene B_4; CRP: C-reactive protein; HNE: human neutrophil elastase; MMPs: matrix metalloproteinases; α_1-AT: α_1-antitrypsin; SLPI: secretory leukocyte protease inhibitor; TIMPs: tissue inhibitors of matrix metalloproteinases; EGF: epidermal growth factor; $D_{L,CO}$: diffusing capacity of the lung for carbon monoxide; CT: computed tomography; PET: positron emission tomography; MR: magnetic resonance; MRC: Medical Research Council; BDI/TDI: baseline and transition dyspnoea index; UCSD: University of California, San Diego; CRDQ: Chronic Respiratory Disease Questionnaire; SGRQ: St. George's Respiratory Questionnaire; BPQ: Breathing Problems Questionnaire; PFSDQ: Pulmonary Function Status & Dyspnoea Questionnaire; PFSS: Pulmonary Function Status Scale; CCQ: Clinical COPD Questionnaire; SF-36: Short Form-36; NHP: Nottingham Health Profile; EQ-5D: EuroQol 5D.

airways or sampled from the lungs. There is recent evidence that blood levels of the acute phase protein C-reactive protein (CRP) may be a marker of inflammation in the airways, and respond to inhaled therapy [49, 50].

In the context of biological markers, a distinction should be made between those use for assessing disease activity or severity in the stable state and those used to define an acute exacerbation (which, by analogy, would be the COPD equivalent to troponin tests for acute myocardial infarction). The identification of biomarkers for exacerbations provides a good example of one of the key problems in developing a biomarker: the need for a precise definition of the disease phenotype. Currently there is no agreed method of identifying an acute exacerbation in clinical terms. This makes it very difficult to identify a precise, sensitive and specific biomarker. It is possible that the reverse may occur, eventually. Basic science may identify a biomarker of the acute inflammation that characterises an exacerbation, thereby providing a method of validating simple clinical methods for identifying when one occurs in routine practice. An exacerbation of COPD presents an interesting discussion point in the context of markers and outcomes. It is probably both an outcome with important consequences for the patient and a marker of underlying chronic disease processes.

Even if some of the putative biomarkers for COPD do demonstrate the required sensitivity and specificity, their introduction to routine clinical practice will require a further, substantial investment of time and money to support the development of assays, logistics and infrastructure. A further problem is that validation of many of these markers will require tests that are not currently performed routinely in clinical laboratories. To date, most studies that contribute to the validation of this kind of marker have been relatively small.

Physiological markers
Identification and validation of physiological markers is not straightforward. Unlike biomarkers, which generally represent a well-defined chemical entity, physiological markers often require agreement between investigators over definitions and methods of measurement. Bronchial hyperreactivity is one example for which there is still no fully standardised approach after nearly three decades [51, 52]. Measures of small airways obstruction held the potential to provide a more comprehensive analysis of the morbidity associated with COPD [53, 54], but this has never been realised. More recently, inspiratory vital capacity [55, 56] and dynamic hyperinflation [57, 58] have been identified as being important markers of COPD; both are currently difficult to perform and standardise. Inspiratory vital capacity may be incorporated into clinical trials, especially

those involved with pharmacological therapies, and reach routine practice, but dynamic hyperinflation will remain a sophisticated laboratory test.

Physiological markers may change over time, and this has led to the development of derived parameters such as decline in FEV_1 over 1 yr. This has become a well-accepted marker of deterioration in COPD that can change with therapeutic intervention, such as smoking cessation [22], but it requires multiple careful measurements made at intervals over a number of years and is more suitable for groups of patients than individuals.

Arterial oxygen tension (Pa,O_2) is another example of an important physiological marker that is measured as a continuous variable, but used to categorise hypoxia into severe or less severe, depending upon whether the Pa,O_2 lies below or above a specific cut-off value (7.5 kPa).

Markers of structure and functional anatomy

Until recently, quantitative measurement of lung structure has required lung resection or *post mortem* specimens, but methods of standardising lung imaging using high-resolution computed tomography scanning are becoming established [59–61]. This is enabling multicentre trials to use this type of measurement as a surrogate marker for changes in lung structure due to COPD. More recently, functional imaging through positron emission tomography and magnetic resonance imaging using hyperpolarised helium and xenon are promising new developments [62–64], although standardisation is still far off.

Symptomatic markers

Many aspects of COPD can still only be assessed through patients reporting symptoms. Some of these might just involve the straightforward recording of symptoms, such as cough, wheeze, breathlessness and sputum colour. The latter provides an example of a surrogate marker, since there is evidence that green-coloured sputum is associated with a higher likelihood of the presence of a bacterial infection in COPD than mucoid sputum [65]. When recording symptoms, standardisation is important, but there is often no agreed form of words or scaling system for quantifying symptom data. This is important because in asthma it has been shown that the wording of simple global measures of severity will affect patients' responses [66]. Furthermore, with few exceptions, there is no agreement over the response options. One notable area of exception is scaling of dyspnoea. There are numerous scales for this, and many of them are validated and standardised [67–69], but there are no methods of converting between them so it is not possible to make direct comparisons of breathlessness measurements made using different scales and units.

Issues concerning marker development become more complex when different components are combined to produce a single overall measurement. This may be done to produce a summary score for a range of symptoms or to produce a scale for a theoretical construct, such as health status. Numerous questionnaires have been developed for measuring impaired health in COPD, reviewed recently elsewhere [70]. As with the FEV_1, it is now possible to calculate the annual rate of deterioration in health for some health status questionnaires, and show that this changes with treatment [71]. Unlike many COPD markers,

indicative values for the threshold of clinical significance are available for some health status questionnaires [72]. Validation of these questionnaires is a complex task and, until recently, it has not been possible to fully test the measurement properties of these instruments. The introduction of probabilistic models, such as the Rasch 1-parameter model, now enables testing of the fundamental measurement properties of questionnaires [73]. This will allow existing questionnaires to be modified or new questionnaires to be created with measurement properties akin to those for physical, physiological and biological measurements.

One specific area in which more work is needed concerns the establishment of equivalence between different questionnaires that purport to measure the same thing. In one respect that is already possible, since the minimum clinically important difference (MCID) has been established for some instruments. The MCID can then be used as a reference point against which to judge the size of change with treatment. A discussion of MCIDs is beyond the scope of this article, but issues of MCID across a wide range of COPD markers have recently been addressed in an entire journal volume [74].

Extrapulmonary markers

With the recognition that extrapulmonary aspects of COPD are important, standardised markers of skeletal muscle function and lean body mass are needed [75, 76]. Exercise capacity is an important marker that results from a range of effects of COPD. Field tests of exercise capacity, such as the 6-min walking distance test, require standardisation, particularly in encouragement [77], but normal values have been reported [78] and a threshold for clinical significance established [79]. The shuttle walking test is standardised [80] but is not yet documented as widely as the 6-min walk. Laboratory-based exercise tests, such as the anaerobic threshold and the symptom limited maximum exercise capacity, are measures of global (*i.e.* pulmonary plus extrapulmonary) consequences of COPD [81, 82]. There is, as yet, no internationally agreed method of defining these and measuring these parameters.

Biological markers of inflammation may also be extrapulmonary markers. For example, there is an association between increased plasma CRP level and impaired lung function in COPD [49]. However, it is important to clarify the nature of such associations. Issues such as whether a systemic marker of inflammation is a ''spill-over'' from a site primarily in the lungs, a marker of a systemic inflammatory process that is causing damage in the lungs, or the result of secondary process in other organs resulting from primary disease in the lungs, will have to be resolved before reliable inferences can be drawn from the use of such markers.

Composite markers

COPD is a complex, multifaceted disease, and reference has already been made to summary measures such as exercise capacity and health status measurements. These markers produce a score that reflects a range of effects of the disease, but another recent approach has been to create a composite made up of markers known to be predictors of mortality: BMI, FEV_1, dyspnoea and exercise capacity (the BODE index) [83]. This has proved to be a better predictor of mortality than the FEV_1 alone. As with all validated composite indices, each

composite measurement must be made in the same way as during the validation studies for the global marker to be reliable.

Properties of markers and interpretation of their measurements

Markers are used to measure different aspects of disease and provide clinically useful information, so it is important to clarify the difference between the properties of the marker and the clinical inferences to be drawn from measurements made with it.

The properties required of a marker will depend on the context in which it is used. For diagnostic markers, there should be a clear distinction between the normal and abnormal range. The ideal diagnostic marker would have a nonlinear performance characteristic, changing rapidly at the transition from normal to the pathological state. More frequently that is not the case and the boundary is a defined point along a continuum, the FEV_1/forced vital capacity (FVC) ratio needed for a diagnosis of COPD is a case in point. Validation of these markers requires some specific steps. First, the range of normality must be ascertained using population-based studies and appropriate statistical methods. Next it is necessary to identify criteria for the presence of the disease, which may be a problem because if there were already one reliable criterion there may be no need for another, unless it was simpler and cheaper to use. For that reason it is often necessary to have multiple criteria. For example, to validate a substance found in exhaled breath condensate as a diagnostic marker of COPD, the minimum criteria for the presence of the disease would probably include a clinical history characteristic of COPD, significant smoking history and a post-bronchodilator FEV_1/FVC ratio <70%. Having established a possible threshold value of the marker that indicates the presence of the disease, it is then necessary to test its sensitivity and specificity in other patients and other settings.

By contrast, markers of efficacy should be sensitive to small treatment effects across a broad range of disease severity. This requires a linear performance characteristic, ideally with finely graded measurements to allow precision. It is usually assumed (but rarely proven) that markers have true interval scaling properties, *i.e.* they behave like rulers. With an interval scale, the distance between two points on the scale should mean the same thing, regardless of the position along the ruler that the two points are placed at. Reliable measurement properties of this type are very important; irregular or nonlinear scaling can cause major problems in data interpretation.

Whilst it is important to ensure that the distance between points on a scale is consistent, this does not necessarily mean that the implication of a given difference in level of disease marker is the same at all points along the scale. An analogy with temperature may be useful. Thermometers have interval-scaling properties, but there are important and different implications for the physical properties of water of a change in temperature between -1 and $+1^\circ$C and between +99 and $+101^\circ$C. Similarly in COPD, the impact on outcomes of a 200 mL change in FEV_1 for a subject whose FEV_1 is 3.0 L will be very different from that for a patient with an FEV_1 of 700 mL. It is very important to make a clear distinction between change in the marker and the clinical implication of that change, *i.e.* its effect on outcome. In many cases an equivalent change in marker will produce very different changes in outcome, often dependent upon the patient's baseline state.

Conclusion

The range of markers available for the assessment of chronic obstructive pulmonary disease and comparison of the effectiveness of different management strategies is currently rather limited. Whilst the forced expiratory volume in one second has come to be used almost as a global marker of chronic obstructive pulmonary disease, it does not reflect the full burden of chronic obstructive pulmonary disease on patients or the multicomponent nature of the disease. Consequently, identification and validation of new markers is needed to provide further insights into the pathogenesis and epidemiology of chronic obstructive pulmonary disease, as well as to allow a more comprehensive assessment of new therapeutic interventions. While the development and validation of markers is a difficult task requiring considerable time and resources, progress in the area is critical to improved understanding and management of this chronic, debilitating and increasingly common disorder.

ACKNOWLEDGEMENTS

REFERENCES

1 Nishimura K, Izumi T, Tsukino M, Oga T. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. *Chest* 2002; 121: 1434–1440.

2 Agusti AG, Noguera A, Sauleda J, Sala E, Pons J, Busquets X. Systemic effects of chronic obstructive pulmonary disease. *Eur Respir J* 2003; 21: 347–360.

3 Colburn WA. Selecting and validating biologic markers for drug development. *J Clin Pharmacol* 1997; 37: 355–362.

4 Rolan P. The contribution of clinical pharmacology surrogates and models to drug development – a critical appraisal. *Br J Clin Pharmacol* 1997; 44: 219–225.

5 Rolan PE. General characteristics of clinical surrogates. *In*: Olesen J, Moskowitz M, eds. Advances in Headache Research. New York, Lippincott-Raven, 1995; pp. 27–32.

6 Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* 1999; 282: 790–795.

7 Cesari M, Penninx BW, Newman AB, *et al.* Inflammatory markers and cardiovascular disease (The Health, Aging and Body Composition [Health ABC] Study). *Am J Cardiol* 2003; 92: 522–528.

8 Kuller LH, Tracy RP, Shaten J, Meilahn EN. Relation of C-reactive protein and coronary heart disease in the MRFIT nested case-control study. Multiple Risk Factor Intervention Trial. *Am J Epidemiol* 1996; 144: 537–547.

9 Smaldone GC, Foster WM, O'Riordan TG, Messina MS, Perry RJ, Langenback EG. Regional impairment of mucociliary clearance in chronic obstructive pulmonary disease. *Chest* 1993; 103: 1390–1396.

10 Jeffery PK. Remodeling in asthma and chronic obstructive lung disease. *Am J Respir Crit Care Med* 2001; 164: S28–S38.

11 Jeffery PK. Comparison of the structural and inflammatory features of COPD and asthma. Giles F. Filley Lecture. *Chest* 2000; 117: Suppl. 1, 251S–260S.

12 Agusti AG, Sauleda J, Miralles C, *et al.* Skeletal muscle apoptosis and weight loss in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002; 166: 485–489.

13 Siafakas NM, Vermeire P, Pride NB, *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.

14 American Thoracic Society/European Respiratory Society Task Force. Standards for the Diagnosis and Management of patients with COPD [Internet]. Version 1.2. New York, American Thoracic Society; 2004. www-test.thoracic.org/copd. Date last updated: September 8, 2005. Date last accessed: September 13, 2005.

15 Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management and prevention of COPD. NHLBI/WHO Workshop Report 2005. www.goldcopd.com/Guidelineitem.asp?11=2&12=1&intId=989. Date last updated: September, 2005. Date last accessed: September 13, 2005.

16 Celli BR. ERS/ATS COPD guidelines: novel aspects. Diagnosis and staging. 13th Annual Congress of the European Respiratory Society. Vienna, Austria, September 27–October 1, 2003; Abstract 2423.

17 Committee for Proprietary Medicinal Products (CPMP). Points to consider on clinical investigation of medicinal products in the chronic treatment of patients with chronic obstructive pulmonary disease (COPD). European Agency for the Evaluation of Medicinal Products (562/98). www.emea.eu.int/pdfs/human/ewp/056298.pdf. Date last updated: October 21, 2005. Date last accessed: April, 2004.

18 Ries AL, Kaplan RM, Limberg TM, Prewitt LM. Effects of pulmonary rehabilitation on physiologic and psychosocial outcomes in patients with chronic obstructive pulmonary disease. *Ann Intern Med* 1995; 122: 823–832.

19 Lacasse Y, Wong E, Guyatt GH, King D, Cook DJ, Goldstein RS. Meta-analysis of respiratory rehabilitation in chronic obstructive pulmonary disease. *Lancet* 1996; 348: 1115–1119.

20 Griffiths TL, Burr ML, Campbell IA, *et al.* Results at 1 year of outpatient multidisciplinary pulmonary rehabilitation: a randomised controlled trial. *Lancet* 2000; 355: 362–368.

21 Celli BR. The importance of spirometry in COPD and asthma: effect on approach to management. *Chest* 2000; 117: Suppl. 2, 15S–19S.

22 Anthonisen NR, Connett JE, Kiley JP, *et al.* Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV_1. The Lung Health Study. *JAMA* 1994; 272: 1497–1505.

23 Vincken W, van Noord JA, Greefhorst AP, *et al.* Improved health outcomes in patients with COPD during 1 yr's treatment with tiotropium. *Eur Respir J* 2002; 19: 209–216.

24 Casaburi R, Mahler DA, Jones PW, *et al.* A long-term evaluation of once-daily inhaled tiotropium in chronic obstructive pulmonary disease. *Eur Respir J* 2002; 19: 217–224.

25 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.

26 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: 449–456.

27 Calverley PM, Boonsawat W, Cseke Z, Zhong N, Peterson S, Olsson H. Maintenance therapy with budesonide and formoterol in chronic obstructive pulmonary disease. *Eur Respir J* 2003; 22: 912–919.

28 Hajiro T, Nishimura K, Tsukino M, Ikeda A, Oga T. Stages of disease severity and factors that affect the health status of patients with chronic obstructive pulmonary disease. *Respir Med* 2000; 94: 841–846.

29 Schols AM, Slangen J, Volovics L, Wouters EF. Weight loss is a reversible factor in the prognosis of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 1791–1797.

30 Landbo C, Prescott E, Lange P, Vestbo J, Almdal TP. Prognostic value of nutritional status in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 160: 1856–1861.

31 Oga T, Nishimura K, Tsukino M, Sato S, Hajiro T. Analysis of the factors related to mortality in chronic obstructive pulmonary disease: role of exercise capacity and health status. *Am J Respir Crit Care Med* 2003; 167: 544–549.

32 Gerardi DA, Lovett L, Benoit-Connors ML, Reardon JZ, ZuWallack RL. Variables related to increased mortality following out-patient pulmonary rehabilitation. *Eur Respir J* 1996; 9: 431–435.

33 Bowen JB, Votto JJ, Thrall RS, *et al.* Functional status and survival following pulmonary rehabilitation. *Chest* 2000; 118: 697–703.

34 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: 680–685.

35 Mildvan D, Landay A, De Gruttola V, Machado SG, Kagan J. An approach to the validation of markers for use in AIDS clinical trials. *Clin Infect Dis* 1997; 24: 764–774.

36 Keatings VM, Barnes PJ. Granulocyte activation markers in induced sputum: comparison between chronic obstructive pulmonary disease, asthma, and normal subjects. *Am J Respir Crit Care Med* 1997; 155: 449–453.

37 Woolhouse IS, Bayley DL, Stockley RA. Effect of sputum processing with dithiothreitol on the detection of inflammatory mediators in chronic bronchitis and bronchiectasis. *Thorax* 2002; 57: 667–671.

38 Yamamoto C, Yoneda T, Yoshikawa M, *et al.* Airway inflammation in COPD assessed by sputum levels of interleukin-8. *Chest* 1997; 112: 505–510.

39 Keatings VM, Collins PD, Scott DM, Barnes PJ. Differences in interleukin-8 and tumor necrosis factor-alpha in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996; 153: 530–534.

40 Crooks SW, Bayley DL, Hill SL, Stockley RA. Bronchial inflammation in acute bacterial exacerbations of chronic bronchitis: the role of leukotriene B4. *Eur Respir J* 2000; 15: 274–280.

41 Woolhouse IS, Bayley DL, Stockley RA. Sputum chemotactic activity in chronic obstructive pulmonary disease: effect of alpha(1)-antitrypsin deficiency and the role of leukotriene B(4) and interleukin 8. *Thorax* 2002; 57: 709–714.

42 Costabel U, Bross KJ, Reuter C, Ruhle KH, Matthys H. Alterations in immunoregulatory T-cell subsets in cigarette smokers. A phenotypic analysis of bronchoalveolar and blood lymphocytes. *Chest* 1986; 90: 39–44.

43 Niewoehner DE, Kleinerman J, Rice DB. Pathologic changes in the peripheral airways of young cigarette smokers. *N Engl J Med* 1974; 291: 755–758.

44 Rutgers SR, Postma DS, ten Hacken NH, *et al.* Ongoing airway inflammation in patients with COPD who do not currently smoke. *Thorax* 2000; 55: 12–18.

45 Retamales I, Elliott WM, Meshi B, *et al.* Amplification of inflammation in emphysema and its association with latent adenoviral infection. *Am J Respir Crit Care Med* 2001; 164: 469–473.

46 Saetta M. Airway inflammation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 160: S17–S20.

47 Saetta M, Di Stefano A, Turato G, *et al.* CD8+ T-lymphocytes in peripheral airways of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 822–826.

48 Fabbri LM, Romagnoli M, Corbetta L, *et al.* Differences in airway inflammation in patients with fixed airflow obstruction due to asthma or chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2003; 167: 418–424.

49 Gan WQ, Man SF, Senthilselvan A, Sin DD. Association between chronic obstructive pulmonary disease and systemic inflammation: a systematic review and a meta-analysis. *Thorax* 2004; 59: 574–580.

50 Sin DD, Lacy P, York E, Mann SFP. Effects of fluticasone on systemic markers of inflammation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2004; 170: 760–765.

51 Chinn S. Comparing and combining studies of bronchial responsiveness. *Thorax* 2002; 57: 393–395.

52 Postma DS, Kerstjens HA. Characteristics of airway hyperresponsiveness in asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 158: S187–S192.

53 Shaw RJ, Djukanovic R, Tashkin DP, Millar AB, du Bois RM, Orr PA. The role of small airways in lung disease. *Respir Med* 2002; 96: 67–80.

54 Hayes GB, Christiani DC. Measures of small airways disease as predictors of chronic obstructive pulmonary disease. *Occup Med* 1993; 8: 375–395.

55 Molho M, Shulimzon T, Benzaray S, Katz I. Importance of inspiratory load in the assessment of severity of airways obstruction and its correlation with CO_2 retention in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1993; 147: 45–49.

56 Taube C, Kanniess F, Gronke L, *et al.* Reproducibility of forced inspiratory and expiratory volumes after bronchodilation in patients with COPD or asthma. *Respir Med* 2003; 97: 568–577.

57 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.

58 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.

59 Bankier AA, Madani A, Gevenois PA. CT quantification of pulmonary emphysema: assessment of lung structure and function. *Crit Rev Comput Tomogr* 2002; 43: 399–417.

60 Spiropoulos K, Trakada G, Kalamboka D, *et al.* Can high resolution computed tomography predict lung function in patients with chronic obstructive pulmonary disease? *Lung* 2003; 181: 169–181.

61 Leader JK, Zheng B, Rogers RM, *et al.* Automated lung segmentation in X-ray computed tomography: development and evaluation of a heuristic threshold-based scheme. *Acad Radiol* 2003; 10: 1224–1236.

62 Jones HA, Marino PS, Shakur BH, Morrell NW. *In vivo* assessment of lung inflammatory cell activity in patients with COPD and asthma. *Eur Respir J* 2003; 21: 567–573.

63 Wild JM, Paley MN, Kasuboski L, *et al.* Dynamic radial projection MRI of inhaled hyperpolarized ^{3}He gas. *Magn Reson Med* 2003; 49: 991–997.

64 Suga K, Tsukuda T, Awaya H, Matsunaga N, Sugi K, Esato K. Interactions of regional respiratory mechanics and pulmonary ventilatory impairment in pulmonary emphysema: assessment with dynamic MRI and xenon-133 single-photon emission CT. *Chest* 2000; 117: 1646–1655.

65 Stockley RA, O'Brien C, Pye A, Hill SL. Relationship of sputum color to nature and outpatient management of acute exacerbations of COPD. *Chest* 2000; 117: 1638–1645.

66 Barley EA, Jones PW. A comparison of global questions *versus* health status questionnaires as measures of the severity and impact of asthma. *Eur Respir J* 1999; 14: 591–596.

67 Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984; 85: 751–758.

68 Mahler DA, Rosiello RA, Harver A, Lentine T, McGovern JF, Daubenspeck JA. Comparison of clinical dyspnea ratings and psychophysical measurements of respiratory sensation in obstructive airway disease. *Am Rev Respir Dis* 1987; 135: 1229–1233.

69 Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. University of California, San Diego. *Chest* 1998; 113: 619–624.

70 Jones PW. Health status measurement in chronic obstructive pulmonary disease. *Thorax* 2001; 56: 880–887.

71 Spencer S, Calverley PM, Sherwood BP, Jones PW. Health status deterioration in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001; 163: 122–128.

72 Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J* 2002; 19: 398–404.

73 Andrich D. Rasch Models for Measurement. London, Sage Publications, 1998.

74 J COPD 2005; 2: 1–459.

75 Mostert R, Goris A, Weling-Scheepers C, Wouters EF, Schols AM. Tissue depletion and health related quality of life in patients with chronic obstructive pulmonary disease. *Respir Med* 2000; 94: 859–867.

76 Wouters EF, Creutzberg EC, Schols AM. Systemic effects in COPD. *Chest* 2002; 121: 127S–130S.

77 Guyatt GH, Pugsley SO, Sullivan MJ, *et al.* Effect of encouragement on walking test performance. *Thorax* 1984; 39: 818–822.

78 Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. *Eur Respir J* 1999; 14: 270–274.

79 Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. *Am J Respir Crit Care Med* 1997; 155: 1278–1282.

80 Dyer CA, Singh SJ, Stockley RA, Sinclair AJ, Hill SL. The incremental shuttle walking test in elderly people with chronic airflow limitation. *Thorax* 2002; 57: 34–38.

81 Patessio A, Casaburi R, Carone M, Appendini L, Donner CF, Wasserman K. Comparison of gas exchange, lactate, and lactic acidosis thresholds in patients with chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1993; 148: 622–626.

82 Casaburi R. Skeletal muscle dysfunction in chronic obstructive pulmonary disease. *Med Sci Sports Exerc* 2001; 33: S662–S670.

83 Celli BR, Cote CG, Marin JM, *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.