Background: The assessment of asthma control is pivotal to the evaluation of treatment response in individuals and in clinical trials. Previously, asthma control, severity, and exacerbations were defined and assessed in many different ways.

Purpose: The Task Force was established to provide recommendations about standardization of outcomes relating to asthma control, severity, and exacerbations.
severity, and exacerbations in clinical trials and clinical practice, for adults and children aged 6 years or older.

Methods: A narrative literature review was conducted to evaluate the measurement properties and strengths/weaknesses of outcome measures relevant to asthma control and exacerbations. The review focused on diary variables, physiologic measurements, composite scores, biomarkers, quality of life questionnaires, and indirect measures.

Results: The Task Force developed new definitions for asthma control, severity, and exacerbations, based on current treatment principles and clinical and research relevance. In view of current knowledge about the multiple domains of asthma and asthma control, no single outcome measure can adequately assess asthma control. Its assessment in clinical trials and in clinical practice should include components relevant to both of the goals of asthma treatment, namely achievement of best possible clinical control and reduction of future risk of adverse outcomes. Recommendations are provided for the assessment of asthma control in clinical trials and clinical practice, both at baseline and in the assessment of treatment response.

Conclusions: The Task Force recommendations provide a basis for a multicomponent assessment of asthma by clinicians, researchers, and other relevant groups in the design, conduct, and evaluation of clinical trials, and in clinical practice.

Keywords: asthma control; asthma exacerbations; asthma severity; clinical trials; outcome assessment (health care); predictive value of tests

INTRODUCTION

Asthma is a heterogeneous condition. Its natural history includes acute episodic deterioration (exacerbations) against a background of chronic persistent inflammation and/or structural changes that may be associated with persistent symptoms and reduced lung function. Trigger factor exposure combines with the underlying phenotype, the degree of hyperresponsiveness and of airflow obstruction, and the severity of airway inflammation to cause wide variability in the manifestations of asthma in individual patients. The challenge to clinicians and researchers is to quantify such profiles both individually and collectively in such a way as to make the assessment of interventions or comparisons between different populations meaningful.

Since the early days of practice guidelines for asthma (1, 2), the aim of treatment has been to minimize symptoms, optimize lung function, and prevent exacerbations. While referring to this aim, lung function was frequently the primary endpoint. With later recognition of the importance of the patient perspective, and of the poor correlation between lung function and inflammation and symptoms (3, 4), clinical trials and clinical practice have increasingly focused on the assessment of “asthma control.” This is a summary term that implies a global assessment of symptoms, reliever use, lung function, and the frequency/severity of exacerbations. To date, there has been no clear definition of asthma control, and the criteria used in its assessment have varied widely from study to study. This substantially limits the extent to which clinical trial data can be pooled for meta-analysis. The definition of “exacerbation” has also varied within guidelines and between studies. The term is variously used to refer to episodes or events occurring multiple times a week (5), or to severe events requiring hospitalization. “Asthma severity” has also been used to describe either a patient’s overall clinical status, or the intensity of asthma symptoms or exacerbations (6).

The present Task Force was established in response to a symposium at the European Respiratory Society (ERS) Congress in Vienna in September 2003. It was approved initially by the ERS and later extended to include the American Thoracic Society (ATS), the aim being to provide recommendations about the assessment of asthma in clinical trials and clinical practice.

Aims of the Task Force

The primary aim of the Task Force was:

To provide consensus recommendations on standardized definitions and data collection methods for assessing asthma control, asthma severity, and asthma exacerbations in future clinical trials.

The secondary aims were:

1. To provide consensus recommendations on standardized measures of asthma control and exacerbations that can be obtained retrospectively from existing clinical trial data, to maximize the potential for pooling of data, and making comparisons between clinical trials.

2. To provide consensus recommendations on the assessment of asthma control, asthma severity, and asthma exacerbations in clinical practice.

Membership of the Task Force

The Task Force membership was intended to represent a broad spectrum of clinical expertise and clinical trial experience. The Food and Drug Administration (FDA) and European Medicines Evaluation Authority (EMEA) each provided an observer to the Task Force. Written submissions were invited from major respiratory pharmaceutical companies; three submissions were received and circulated to the Task Force members for their information. Employees of pharmaceutical companies and representatives of other companies with a commercial interest in the output of the Task Force were excluded from membership or observer status.

Seven Working Groups of three to five members provided detailed reports about: exacerbations, diary data, physiologic measures, composite scores, biomarkers, indirect measures, and quality-of-life questionnaires. The Working Group reports and the combined report were circulated to all Task Force members for comment and agreement. Pediatric and Primary Care Working Groups provided specialized perspective on the recommendations from each Working Group (see Working Group Membership).

Scope of the Task Force Work

The Task Force considered outcome variables from the perspective of their relevance to clinical trials, including those conducted in primary care, and to clinical practice at all levels from primary care to tertiary care. The focus was primarily on issues relating to the assessment of asthma in those aged 6 years and over, as the needs for children younger than 6 years were considered to be beyond the scope of the present initiative.

METHODOLOGY

Definitions of Asthma Control and Asthma Severity

During preparation of the Task Force proposal in 2004, it was already obvious that there were no consistently accepted definitions for asthma control, severity, or exacerbations. This was subsequently confirmed by literature reviews performed by individual Working Groups (described below). Criteria based on American, British, and international (GINA) guidelines were cited in some clinical trial reports as measures of treatment effect on “asthma control,” and in others as relating to “asthma severity.” The Task Force therefore set about establishing new definitions for asthma control and asthma severity, based on consensus and clinical relevance. After initial round-
Literature Review

The Task Force members agreed that the output from the Task Force would be a narrative review, identifying and describing measures that were appropriate to the newly established definitions of asthma control and exacerbations. It was not appropriate in most cases to grade the quality of studies, as is usually a priority for reviews about clinical efficacy (7).

The Cochrane Register of Randomized Controlled Trials was searched for all studies published between 1998 and 2004 that contained the words “asthma control,” “asthma severity,” or “asthma exacerbation(s).” This yielded 440 references, decreasing to 356 after exclusion of non–English-language references and those published only as abstracts, and to 327 after exclusion of duplicates. Papers reporting studies based only within a laboratory, emergency room (ER), or hospital were also excluded. Thereafter, each paper was allocated to an individual Task Force member to identify the outcome variables that had been used to describe and/or quantify asthma control, severity, and/or exacerbations. This search provided a subset of clinical trial reports that had used outcome variables relevant to previous definitions of asthma control, severity, or exacerbations.

Evaluation of Outcome Measures

Papers identified by the above process were allocated to the relevant Working Group(s) for joint assessment of the measurement properties, analysis and reporting of the outcome variables, and the information that they provided about treatment response. The Working Groups then performed descriptive reviews of each identified outcome variable, based on a customized template (see the online supplement). The focus was on the measurement properties of each variable, and its strengths and weaknesses in the assessment of asthma control or exacerbations (as defined by the Task Force). Task Force members were asked to refer where possible to published guidelines or recommendations about methodology relevant to their Working Group, and to identify additional papers that evaluated measurement properties or clinical associations for the nominated outcome variables from further literature searches. References were updated before the final submission of this document.

Development of Final Recommendations

Finally, after considering the available outcome measures, the Task Force developed overall recommendations for the assessment of asthma control and exacerbations in clinical trials and clinical practice. The recommendations are based on a balance between:

1. The extent to which each measure provides information that is congruous with the definitions of asthma and of asthma control.

2. The extent to which each measure is reflective of the dual goals of asthma treatment (6, 8, 9), namely:
   a. to achieve good control of the current clinical manifestations of asthma, and
   b. to reduce risk to the patient (i.e., the risk of adverse outcomes such as exacerbations, poor control, accelerated decline in lung function, and side-effects of treatment). Some of these future risks may result from lack of control of the underlying disease process.

3. Characteristics of the outcome measures such as reproducibility, responsiveness, and construct validity (association with other measures), all of which will obviously vary to some extent between individual outcome variables within each class.

4. Feasibility of using the outcome measure (including safety, accessibility, and cost).

For each of the main groups of outcome measures, key points and recommendations for clinical trials are presented in summary boxes at the end of the relevant sections below, together with recommendations for clinical practice, pediatric issues, and important research questions. The overall recommendations for assessment of asthma control and exacerbations in clinical trials, and the rationale for each recommendation, are presented at the end of this document (see Tables 1 and 2 in SUMMARY AND OVERALL RECOMMENDATIONS). The Task Force recommendations for the assessment of asthma severity were previously published (6), and are summarized below (see ASTHMA SEVERITY).

TASK FORCE DEFINITIONS

A summary of the Task Force definitions of asthma control, severity, and exacerbations is provided below. The rationale for the development of these definitions has been published in the European Respiratory Journal in 2008 (6), and it is important that the present recommendations should be read in conjunction with that paper.

Asthma Exacerbations

In clinical practice, exacerbations are identified as events characterized by a change from the patient’s previous status. This concept should also be applied in clinical trials.

1. Severe asthma exacerbations are defined as events that require urgent action on the part of the patient and physician to prevent a serious outcome, such as hospitalization or death from asthma.

2. Moderate asthma exacerbations are defined as events that are troublesome to the patient, and that prompt a need for a change in treatment, but that are not severe. These events are clinically identified by being outside the patient’s usual range of day-to-day asthma variation.

Although several studies have reported “mild” exacerbations, the Task Force considered that these episodes were only just outside the normal range of variation for the individual patient and that with present methods of analysis, they could not be distinguished from transient loss of asthma control. Hence, no definition of a “mild” exacerbation can be offered (see ASTHMA EXACERBATIONS for more detail).

Asthma Control

Asthma control is defined as the extent to which the various manifestations of asthma have been reduced or removed by treatment. This includes two components:
1. The level of clinical asthma control, which is gauged from features such as symptoms and the extent to which the patient can carry out activities of daily living and achieve optimum quality of life, and

2. The risk of future adverse events including loss of control, exacerbations, accelerated decline in lung function, and side-effects of treatment. More detail about asthma control is provided in General Concepts about Asthma Control.

Asthma Severity

Asthma severity is defined as the difficulty in controlling asthma with treatment. After exclusion of modifiable factors such as poor adherence, smoking, and comorbidities (10), severity largely reflects the required level of treatment and the activity of the underlying disease state during treatment, which may vary depending on the underlying phenotype, environmental factors, and comorbidities (6). There is clinical utility in distinguishing patients with “difficult-to-treat” or severe asthma from those who have “easy-to-treat” or mild asthma.

This represents a change from previously published definitions of asthma severity (8, 9), which was previously defined in terms of the activity of the underlying disease process as represented by clinical features before commencement of treatment. The Task Force considered that current clinical and research usage of “severe asthma” and “mild asthma” overwhelmingly focused on the intensity of treatment required. In addition, there was insufficient research evidence that a patient’s clinical characteristics when untreated could consistently inform future management decisions, or could predict the ease or difficulty of obtaining good asthma control once treatment was commenced, to warrant retaining the previous “off-treatment” definition of severity. Further explanation is provided in the separate publication (6).

**Asthma Exacerbations**

**Background**

Prevention of asthma exacerbations has been identified in all asthma treatment guidelines (e.g., References 8, 9, 11) as an important component of establishing ideal asthma control. It could be argued that exacerbations are the most important outcome, because they constitute the greatest risk to patients, are a cause of anxiety to patients and their families, result in the greatest stress on health care providers, and generate the greatest cost to the health care system (12). Somewhat surprisingly, only in the past 10 years have exacerbations been used as a primary outcome variable in research into the efficacy of drug treatment in asthma (13).

Exacerbations are recognized as a common clinical manifestation in patients with severe asthma, and are known to increase the risk of asthma mortality (14). However, even in patients thought to have mild asthma, the rates of severe asthma exacerbations have been much higher than expected (15, 16).

In clinical practice, exacerbations are recognized as episodes that are troublesome to patients, and that prompt a need for a change in treatment. These episodes vary considerably in speed of onset, from minutes or hours (17, 18) to 2 weeks (19), and in time to resolution (5–14 d [19]); they also vary in their absolute severity, both between and within individual patients. Clinical characteristics that cause acute distress and impairment in one patient may represent another patient’s usual status. These events are therefore clinically identified by being outside the patient’s own usual range of day-to-day variation.

**Previous Definitions of Exacerbations**

Various terms are used to refer to exacerbations, and this impacts on the yield of literature searches. For example, ER studies often refer to “acute severe asthma” rather than “exacerbations.” Likewise, in ER studies, “mild acute asthma” may refer to episodes with FEV₁ > 30% predicted (20), events that would be regarded as extremely severe in a community-based setting. Some studies refer to “treatment failure” (21), or use “asthma attack” to differentiate a severe exacerbation from an “exacerbation day” (22–24).

**Severe exacerbations.** In the reviewed literature, the definitions of severe exacerbations most frequently included the need for the administration of systemic corticosteroids (tablets, suspension, or injections) (13, 16, 25, 26) at the physician’s discretion (27–29), and/or in response to a specified decrease in peak expiratory flow (PEF) (13, 16, 30). Other criteria variously included emergency room visits and/or hospitalizations (13, 16) or unscheduled doctor visits (30). Some studies excluded use of systemic corticosteroids for less than 3 days (31, 32). Few studies reported whether closely consecutive courses of systemic corticosteroids were handled as one or two exacerbations.

Many definitions included a decline in PEF of 20 to 30% (13, 16, 25, 26, 33). Usually, the criterion was a change from baseline PEF, rather than change from on-treatment PEF. Most studies required at least 2 consecutive days of lower PEF, but occasionally a single day with low PEF was accepted (26). In two studies (13, 33) in which severe exacerbations were identified retrospectively from systemic corticosteroid use and/or changes in PEF, the majority were identified by corticosteroid use. Finally, some studies included criteria for increased asthma symptoms (30, 34) or increased rescue short-acting β₂-agonist (SABA) plus a fall in PEF (26).

**Moderate exacerbations.** Moderate exacerbations may be considered to be events that require additional treatment to prevent progression to severe exacerbation. Few studies formally defined moderate exacerbations. One study defined moderate exacerbations by use of corticosteroid tablets, and severe exacerbations by in-patient care or ER visits (35). Another defined moderate exacerbations by “extra controller therapy” (oral corticosteroids/inhaled corticosteroids [ICS]/long-acting β₂-agonist [LABA]/theophylline) (28). In studies that examined whether doubling (36, 37) or even greater increases (38) in ICS dose for worsening asthma reduced the chance of progression to severe exacerbation, the index events obviously had to be defined by clinical criteria rather than by the medication change itself. In these studies, the use of more stringent clinical criteria (36) did not increase the probability of progression to severe exacerbation (defined by need for oral corticosteroids).

**Mild exacerbations.** In the reviewed studies, a range of criteria were used to define mild exacerbations—for example, a 15% decrease in morning PEF (26), a 20% decline in clinic FEV₁ (28), and/or an increase in reliever medication use (25, 39, 40). Some studies defined “exacerbation days” either singly (23, 36, 41) or as consecutive days (13), but similar features were used in other studies to identify “poor control days” (16).

Vaquerizo and colleagues observed that even subtle differences in criteria for exacerbation days resulted in large differences in the apparent incidence of these events, and could reduce the apparent efficacy of an intervention (23).
Exacerbations of unspecified severity. Many other studies reported exacerbations without classifying them as to severity, using criteria similar to those described above for severe exacerbations (e.g., References 37, 42, and 43). Some studies used composite criteria of symptoms, \( \beta_2 \)-agonist use, PEF, or FEV\(_1\), or an increase in maintenance therapy (e.g., References 41, 44, and 45), but did not report the proportion of exacerbations identified by each criterion. Some studies required only a single criterion to be identified (e.g., Reference 44), while others needed two or more (e.g., Reference 46).

Utility and Implications of Previous Definitions

Severe exacerbations. Of those evaluated, almost no two studies had the same definition of a severe asthma exacerbation. Most included the need for systemic corticosteroids (or an increase in dose of maintenance oral corticosteroids) and/or hospitalization/ER visit. These criteria appear to be clinically relevant and intuitively valid. However, they are open to the criticism that, as the reasons for the decision to begin corticosteroids or visit the ER are not defined a priori, the events will differ from study to study, patient to patient, and country to country. In a primary care study from The Netherlands, exacerbations defined by oral corticosteroid use had low sensitivity for doctor-diagnosed exacerbations in primary care (47). It should be recognized that both systemic corticosteroid use (initiated by patient or clinician) and ER/hospitalization require a subjective assessment by the patient and/or clinician that the event is severe enough to warrant such action. These definitions are therefore dependent on adequate perception of airway obstruction by the patient. However, at a group level in clinical trials, use of systemic corticosteroids and/or urgent health care utilization have been found to be responsive to treatment (e.g., Reference 33), and they are relatively simple to record.

The Task Force recommends that the definition of severe exacerbations should include at least 3 days’ use of systemic corticosteroids, to avoid including inadvertent or inappropriate patient-initiated use. If systemic corticosteroids are used for less than 3 days, the reason for discontinuation (e.g., side effects), should be recorded. While this 3-day criterion has never been critically evaluated and is likely to suffer from lack of precision, it is clinically relevant, as a shorter duration of treatment is not recommended by guidelines. For closely consecutive courses of corticosteroids, some studies count two courses separated by at least 1 week as separate severe exacerbations; this lacks firm evidence, but there are advantages in a standardized approach.

The definition of severe exacerbation in terms of systemic corticosteroid use or hospitalization is suited only to retrospective use—for example, in the assessment of clinical trial outcomes. It does not provide clinical guidance for diagnosis and management of exacerbations during a clinical trial or in clinical practice.

In several studies, a poor association has been observed between PEF criteria and clinician prescription of corticosteroids. This calls into question the clinical relevance of PEF in defining severe exacerbations. In some studies, many episodes with 30% fall in PEF were identified retrospectively from diaries, without the patients having presented for medical care (13, 33). This may have been due to lack of symptoms or a reluctance to take systemic corticosteroids (as reported in Reference 48); or the relevant paper diary data may have been completed retrospectively (49). By contrast, in other studies (30, 50), a 30% decline in PEF was found to be too stringent a criterion, with the majority of patients initiating extra treatment before this level had been reached. The disparity between the above observations may be explained by differences in baseline PEF variability between patients (51) or to other monitoring-related factors. However, at a group level, exacerbations defined by a greater than 30% decline from baseline for 2 consecutive days are responsive to long-term treatment (13, 33). Data from paper PEF diaries should be interpreted with caution in the analysis of exacerbations, because of the high proportion of retrospectively completed entries (49) (see DIARY DATA IN THE ASSESSMENT OF ASTHMA CONTROL). Further work on diagnostic criteria for exacerbations is needed based on electronically recorded data.

Although increased symptoms and \( \beta_2 \)-agonist use are characteristic of severe exacerbations, it is difficult to establish appropriate criteria for general use, given the range of symptoms and medication use before the exacerbation. Nocturnal symptoms appear to develop late in the course of an exacerbation (30), and therefore may be an insensitive criterion to use for the definition of severe exacerbation. In selecting criteria to prospectively define severe exacerbations, one must balance specificity against safety. Changes that occur for only 1 day may potentially reflect transient loss of asthma control, rather than an exacerbation, but for a patient who is rapidly deteriorating at the beginning of a severe exacerbation, a 2-day wait to initiate additional treatment may be too long.

Moderate and mild exacerbations. The concept of a moderate exacerbation has clinical utility, as clinical practice guidelines advise that exacerbations should be recognized and treated before they become severe. Despite the lack of validated criteria, it appears reasonable for a definition of moderate asthma exacerbations to include deterioration in symptoms and/or lung function with increased rescue bronchodilator use that lasts 2 days or more, but that is not severe enough to warrant corticosteroid tablet use and/or a hospital visit. However, the way in which such deterioration can be standardized for either prospective or retrospective use in clinical trials has not yet been established.

Defining mild exacerbations is difficult, because by definition these episodes will be only just outside the patient’s normal range of variation. Defining the end of an exacerbation, particularly if it represents only poor control, is even more challenging.

Analysis and reporting of exacerbation data. Some studies have analyzed the percentage of patients with at least one exacerbation (e.g., Reference 44) or the time to first severe exacerbation (e.g., Reference 16). The latter has the advantage, particularly in placebo-controlled trials, that it is less likely to be contaminated by the introduction of additional therapies, or affected by patients with multiple exacerbations. However, most commonly, studies have used the annualized rate of exacerbations (13, 16, 46, 52), which has been useful in comparing patient populations in different studies. The statistical issues involved in analysis of exacerbation rates have been described in detail for chronic obstructive pulmonary disease (COPD) (53, 54), and are also applicable in asthma. The weighted mean (total number of exacerbations for the study group, divided by the total duration of person follow-up for the group) provides the best statistical estimate of the exacerbation rate, as it is unbiased by the effect of exacerbations occurring in a small time interval (53).

Exacerbation rates and proportion of subjects with at least one exacerbation have been most commonly displayed by column graphs (e.g., Reference 52). Time to first exacerbation is often displayed by Kaplan-Meier survival graphs (e.g., References 16 and 55), with separate graphs used for time to second or third exacerbation. More recently, composite graphs have been used to display individual exacerbations for each study subject, thus showing the time to occurrence and duration of both initial and repeat exacerbations within the one graph (e.g., References 56 and 57).
Severe Asthma Exacerbations
1. Severe asthma exacerbations are events that require urgent action on the part of the patient and physician to prevent a serious outcome, such as hospitalization or death from asthma. The occurrence of severe asthma exacerbations should be used as a marker of poor asthma control.

2. The definition of a severe asthma exacerbation for clinical trials should include at least one of the following:
   (a) Use of systemic corticosteroids (tablets, suspension, or injection), or an increase from a stable maintenance dose, for at least 3 days. For consistency, courses of corticosteroids separated by 1 week or more should be treated as separate severe exacerbations.
   (b) A hospitalization or ER visit because of asthma, requiring systemic corticosteroids.

3. If severe exacerbations are defined by any of multiple criteria, investigators should be asked to record the specific criterion (or criteria) which were satisfied for each episode, so that they can be better characterized.

4. The inclusion of a percentage change in PEF from baseline is not currently recommended as a criterion for severe exacerbations.

5. There are currently no validated criteria for the magnitude of change in symptoms and/or β2-agonist use that define a severe asthma exacerbation. If included in a study, changes in PEF, symptoms, and/or β2-agonist use should persist for 2 or more days (unless very severe) to qualify as a severe exacerbation.

6. For individual patients, information about the onset and duration of exacerbations should subsequently be used to refine and customize the “trigger points” or “action points” for the patient’s asthma action plan.

EXACERBATIONS IN CLINICAL PRACTICE

1. Exacerbation frequency should be evaluated as part of routine asthma assessment, and so a definition for exacerbations is needed for clinical practice.

2. There is an urgent need for prospective rather than retrospective definitions of exacerbations for clinical practice, to provide guidance for health care professionals and patients in treating exacerbations.

3. Defining exacerbations on the basis of systemic corticosteroid use has low sensitivity for doctor-diagnosed exacerbations in primary care.

4. Prospective definitions that may be suitable for clinical trials are not necessarily suitable for clinical practice, because of differences in health care resources (e.g., 24-h access to investigators in clinical trials), and patient and clinician expectations about, and experience with, monitoring of symptoms or PEF.

5. In the clinical setting, the absolute severity of exacerbations will vary considerably from patient to patient, or over time. Therefore, asthma exacerbations should be clinically identified by changes in symptoms and/or rescue use and/or lung function, which are outside the patient’s usual range of day-to-day asthma variation, and, for retrospective analysis, are associated with an increase from a stable maintenance dose of treatment taken by the patient for 3 days or more.

6. For individual patients, information about the onset and course of exacerbations should subsequently be used to refine and customize the “trigger points” or “action points” for the patient’s asthma action plan.
PEDIATRIC ISSUES

1. In young children with asthma, exacerbations are frequent, with significant morbidity, possibly because of the frequency of viral infections.

2. There are no reliable methods for early detection, but the development of upper airway symptoms of viral infection may be a useful alert.

3. The severity of exacerbations is also difficult to characterize in children, because of dependence on parental reporting of symptoms and the difficulty of measuring lung function.

4. Many exacerbations in children are treated with increased doses of ICS rather than systemic corticosteroids; however, until specific studies are available, these should be considered moderate exacerbations, while the use of systemic corticosteroids would constitute a severe exacerbation.

RESEARCH QUESTIONS

1. More work is needed to establish clinical criteria for changes in symptoms, β₂-agonist use, and lung function that can be used prospectively to identify exacerbations. This analysis should be based on electronically recorded data, and should take into account each patient’s usual range of variation (e.g., in symptoms or PEF).

2. The factors (clinical, psychological, and contextual) that contribute to patient and clinician decisions to use systemic corticosteroids or that prompt urgent health care utilization need to be further investigated.

3. More work is needed to develop clinical criteria for pediatric asthma exacerbations that can be applied prospectively.

4. More work is needed to develop simple feasible criteria for defining exacerbations in clinical practice.

GENERAL CONCEPTS ABOUT ASTHMA CONTROL

The current definition of asthma (5, 8) comprises four domains: symptoms, variable airway obstruction, airway hyperresponsiveness (AHR), and airway inflammation. No one domain is essential to the diagnosis, and not all investigators have access to objective testing of all four domains. In primary care, the diagnosis of asthma is often made only on the basis of symptoms, but, given their lack of specificity, this approach may lead to incorrect diagnosis. Confirmation of the diagnosis of asthma once regular treatment has been commenced is even more difficult. There is increasing awareness in the literature and in clinical practice of the importance of different asthma phenotypes and their differences in responsiveness to treatment (58, 59). Hence, the recommendations of the Task Force about the assessment of asthma control are governed not only by the performance characteristics of the outcome variables themselves (as described in the remainder of this document), but also by the definition of asthma, the concept of differing asthma phenotypes, and the goals of asthma treatment.

Current Clinical Control and Future Risk

Asthma control is defined as the extent to which the various manifestations of asthma are reduced or removed by treatment. Although severe exacerbations are more common with poorly controlled asthma (52), they also occur in patients with otherwise mild (27) or well-controlled asthma (60). In addition, some medications, such as LABA (given alone), may control symptoms and lung function in the short term without reducing inflammation or AHR (41, 61). Hence, there is increasing recognition in asthma guidelines (9) of the need for the concept of asthma control to encompass not only the patient’s recent clinical state (symptoms, night waking, reliever use, and lung function), but also to consider their “future risk”—that is, their potential for experiencing adverse outcomes, such as loss of control in the near or distant future, exacerbations, accelerated decline in lung function, or treatment-related side effects.

While current poor control predicts future poor control and health care utilization (62), there is increasing awareness that other pathologic and physiologic measures, independent of the level of current clinical control, predict future risk. For example, exhaled nitric oxide has been used as a “predictor” of loss of asthma control (63–65). Some independent risk factors may be identified from bioinformatics, that is, the use of mathematical tools to extract useful information from large datasets (e.g., Reference 66). In short-term studies, where long-term risks are not be able to be recorded, some outcome measures can be used as surrogate markers for change in future risk.

Validation of Measures of Asthma Control

Because there is no gold standard for the definition of asthma, there can be no gold standard for the assessment of asthma control, and no single primary endpoint can be recommended for the assessment of treatment response in asthma. The four components of the definition of asthma (symptoms, airway obstruction, airway hyperresponsiveness, and airway inflammation) are only loosely associated (67–70), so no one of these domains is completely suitable as a comparator for validation of individual measures of asthma control.

The comparator for validation of asthma control measures has often been a “global physician assessment,” but this may not provide an absolute standard (71). In the past, global physician assessments were probably based largely on the patient’s recent clinical status (symptoms, night waking, reliever use, exacerbations) as recommended in clinical practice guidelines (8), together with factors such as the extent of patient “bother” (72). The shift in recent guidelines (9) to explicitly incorporate future risk into the assessment of asthma control may lead to a subtle change in physician global assessments in years to come. At present, there are no firm guidelines as to how physicians should integrate the dual components of current control and future risk into an overall assessment of asthma control, either in clinical practice or in clinical trials.

Asthma-related quality of life is a global measure of the impact of asthma from the patient’s perspective and has been used for validation of some of the measures examined in this document. The patient’s own assessment of their level of asthma “control” will depend on their ability to detect airway obstruction (73), as well as their personal interpretation of the term “control.” This may differ markedly from that of the clinician (6).

In asthma, response to treatment is not an absolute characteristic. It depends on the baseline status of the study population (including phenotype), the mechanism of action, dose and duration of treatment, and the properties of the outcome measure itself. A medication may lead to a response in one domain of asthma or one phenotype, and hence in some outcome measures but not others. For example, early clinical trials of anti–interleukin-5 showed a profound reduction in eosinophils but no effect on AHR (74). The insight into pathophysiologic mechanisms that was provided by this and similar studies highlights the benefit of assessing more than one domain of asthma in clinical trials.
Range of Asthma Control

Asthma control is best considered as a scale or continuum; that is, one should not refer to “achieving asthma control” to imply achieving good control (6). In the past, the level of control has often been categorized using semi-quantitative descriptors such as “total,” “good,” or “poor,” or using relative terms ranging from “best achievable” or “optimal” at one end of the continuum, through “sub-optimal” to “undesirable” or “unacceptable” at the other end. However, such descriptors involve arbitrary cut-points, and so continuous or ordinal measures are preferred. If categorical descriptors are needed (e.g., to describe study populations at baseline), they should be based on clinically meaningful cut-points.

Descriptors of asthma control such as “acceptable” beg the question of whose perspective is being considered, the patient’s or the physician’s, and whether the cost (financial and/or drug-related side effects) of achieving the desired level of control has been considered. The primary perspective for the assessment of asthma control is that of the patient. However, the level of control that may be acceptable to one patient may be unacceptable to another or to a clinician. In each case, “acceptability” of the patient’s current clinical state needs to be balanced against the future risk of either poor control or of treatment-related adverse effects, and patient-reported measures should be supplemented by the (objective) measures that relate to the pathophysiologic domains of the definition of asthma.

Time as a Factor in the Assessment of Asthma Control

A patient’s level of asthma control may vary over relatively short time intervals (days to weeks) in response to allergens or infectious agents, or in response to treatment. The time-course of improvement with treatment varies according to the particular outcome variable being measured (Figure 1) and according to the type of treatment (e.g., ICS versus LABA). Therefore, in cross-sectional studies, the relationship between different measures will vary according to the time of assessment. The order of reappearance of different clinical features during an exacerbation or after cessation of ICS also varies (30, 75).

Figure 1. Time-course of improvement in different asthma control outcome variables with inhaled corticosteroid treatment. Reprinted with permission from Reference 418. This figure was constructed with data from Reference 26; the statistical analysis of time to plateau is described in Reference 122. AHR = airway hyperresponsiveness; SABA = short-acting β2 agonist.

In clinical trials, asthma control is assessed at defined points in time, and some variables are suited to such interval measurement. To be clinically relevant, variables such as symptoms must relate to a meaningful recent period (76), rather than just at the moment of assessment. By long-standing consensus, clinical asthma control is usually assessed over periods of 1 to 4 weeks: this approach is supported by empirical data (77).

Applicability of Control Measures to Clinical Trials

Some outcome measures are used to assess whether a particular intervention was active in relation to its specific pharmacologic target. However, for interventions targeted at overall asthma control, it is clearly desirable for outcome measures and surrogate markers to be relevant to more than one therapeutic intervention. Based on experience with antiinflammatory therapy, it is often assumed that future risk of exacerbations will directly parallel changes in current clinical control. However, these two aspects are not necessarily concordant, particularly with LABA monotherapy (41, 61), or even with combination ICS/LABA (78). Such discordance (between current clinical control and future risk) should be considered in the evaluation of other therapies (e.g., a therapy aimed at reducing rhinovirus infections may reduce exacerbations without changing the level of clinical control). Hence, given that the goals of asthma treatment relate to both the achievement of good control and the minimization of future risk, it is not appropriate to specify a single primary endpoint for the assessment of asthma control. Studies of clinical efficacy and effectiveness should use appropriate endpoints which capture both aspects of asthma control.

Analysis of Asthma Control in Clinical Trials

There may be situations in which it is appropriate to assess the “end-of-treatment” level of control (e.g., if the study hypothesis focuses on the proportion of patients who meet a certain criterion after a specified period of treatment). However, in general, it is preferable to assess the level of asthma control throughout the study. Analysis of data drawn from over the whole treatment period will reflect the magnitude and rate of treatment response, the extent of variation in level of control, and the occurrence of exacerbations, all of which are relevant to the overall impact of treatment on asthma control. This will typically include assessment at each study visit, or, for diary measures, over multiple periods each of 1 to 4 weeks. The resulting multiple data points for each patient can be analyzed by mixed model (or equivalent) analyses, which are to be preferred over merely averaging the data over the whole treatment period, to improve the power of the study.

DIARY DATA IN THE ASSESSMENT OF ASTHMA CONTROL

Symptoms and airway obstruction are integral to the definition of asthma, and represent important components of the assessment of asthma control in clinical practice and clinical trials. However, symptoms are highly variable, and interval questionnaires, administered at clinic visits, are limited by patient recall (79) and by the improved medication adherence that is seen in the weeks before and after a clinic visit (80). Airway obstruction is also characteristically variable in asthma, and clinic lung function, if measured after withholding of study medication, does not represent the patient’s usual daily on-treatment state. Therefore, it can be advantageous to use diaries to record outcome variables during the subject’s normal day-to-day life, or to identify the time-course of change with treatment. However, short-term diaries, used for 1 to 2 weeks before clinic visits, may not provide more information about symptoms than can be obtained from clinic-based questionnaires (81). Adherence is a major issue with long-term diaries (49), but high rates of adherence can be achieved with user-friendly electronic diaries (82). Diaries can also be used for recording adverse events, interference with activities, and health care utilization (83, 84).
Methods of Recording Diary Data

Most published diary data have come from paper diaries and mechanical PEF meters. The introduction of electronic devices allowed covert assessment of such records, and the universal finding, in asthma as in other diseases, was that a substantial proportion of data in paper diaries was fabricated or completed retrospectively (49). Electronic devices are now being introduced for routine collection of time-verified diary data in clinical trials (85–87), under draft guidance provided by the FDA. Electronic diaries enhance data quality (88, 89), and can improve adherence with monitoring (82), which should substantially reduce sample size requirements (90). Electronic recording reduces time to database lock (86). These benefits must be weighed against equipment costs, which may be substantial. In addition, careful attention must be paid to reliability of equipment and software, user-friendliness, rapid technical support, regulatory requirements, and practical issues such as after-midnight entries, travel across time zones, and shift work (86, 88, 89). In a meta-analysis, paper- and computer-administered versions of patient-reported outcomes were found to be equivalent (91).

Symptom monitoring. Asthma symptoms are nonspecific, and to be equivalent (91). administered versions of patient-reported outcomes were found as after-midnight entries, travel across time zones, and shift of equipment and software, user-friendliness, rapid technical substantial. In addition, careful attention must be paid to reliability of equipment and software, user-friendliness, rapid technical support, regulatory requirements, and practical issues such as after-midnight entries, travel across time zones, and shift work (86, 88, 89). In a meta-analysis, paper- and computer-administered versions of patient-reported outcomes were found to be equivalent (91).

Diary Questions

Reliever use. Use of quick-acting β₂-agonists may reflect the frequency and intensity of symptoms, the patient’s symptom tolerance, the usual level of physical activity, and the duration of action of any routinely taken β₂-agonist (usually a LABA). In clinical trial reports, β₂-agonist use is usually quantified as the number of inhalations or puffs/day, or as reliever-free days. However, the routine dose of some β₂-agonists is one inhalation, and for others, two inhalations (and even for the latter, some patients may routinely use one inhalation). To reduce this heterogeneity, we recommend that β₂-agonist use should be recorded as “occasions” rather than “puffs” per day. Diary instructions should clearly explain the difference. Patients may interpret “times per day” as meaning either “puffs” or “occasions,” so this wording should also be avoided.

Some β₂-agonist use is anticipatory (e.g., to prevent exercise-induced asthma). This may be reflected in a lower proportion of β₂-agonist–free days than symptom-free days (33, 102). Some studies exclude β₂-agonist use before exercise (103) and others report it separately (104). We recommend that use of β₂-agonist for relief of symptoms should be recorded and reported separately from prophylactic use.

In childhood asthma, rescue medication use is often controlled by the parent (105), and this may produce discrepant reports. In a pediatric study, little relationship was found between electronically recorded SABA use and symptom severity score (106). It is unclear whether supervised rescue medication gives a better assessment of asthma control than medication taken without involvement of a parent. Hence, both symptoms and β₂-agonist use should be reported.

β₂-agonist use provides a continuous numerical measure of asthma control, but reporting presents some challenges. Daily β₂-agonist data are usually right-skewed, but daily medians may conceal intermittent heavy usage (e.g., if a subject uses β₂-agonist on 8 occasions/d for 3 d each week, median usage is still 0 occasions/d). Reporting mean daily usage (total occasions in one week/7) is an improvement, but cannot distinguish between intermittent heavy use and daily low-frequency use. Hence, we recommend reporting of both the proportion of reliever-free days and the mean of occasions per day, the latter averaged over 1 week.

Validation of symptom and reliever diaries. As diaries are a special form of questionnaire, they should be subjected to the same validation procedures as interval questionnaires, but this has rarely been done. Clinical trial outcomes may be dramatically affected by even minor differences in the wording and criteria for diary measures (23), so standardization is important.

Some studies have validated patient diaries against interval clinic questionnaires, but the latter are limited by patient recall. The same problem applies to using physician global assessment as the comparator, as this too is heavily influenced by patient recall of symptoms and reliever use. If electronic diaries are available to avoid recall errors, clinician questionnaires could be validated against daily records rather than vice versa.

Santanello and colleagues have published validation studies for asthma diaries for adults (83), and children aged 6 to 14 years (85). The adult diary showed good internal consistency and reproducibility in two placebo-controlled trials. Longitudinal associations were stronger for average daily PEF than for weekly clinic FEV₁ (83). The daytime pediatric diary had acceptable longitudinal construct validity against physician assessment and quality of life in stable and unstable patients (85). Juniper and colleagues developed and validated a daily diary from the Asthma Control Questionnaire (ACQ), and found that its measurement properties over a single week were similar to those of the ACQ itself, recorded at the end of the same week (82).

Responsiveness. Despite the poor quality of data from paper diaries, diary measures have shown significant treatment responses at a group level in many asthma studies. “Symptom-free days” have generally shown good responsiveness in subjects...
with frequent symptoms at entry (16, 26, 107, 108), but this variable is limited by a “ceiling effect” for patients with mild asthma and a “floor effect” for patients with severe asthma. Responsiveness testing of the Santanello diaries was limited by the limited efficacy of the study medication on the comparator variable (FEV₁) (86). For validation of the ACQ diary, physician assessment was used as the comparator, and hence responsiveness, not surprisingly, favored the clinic questionnaire recorded on the same day as the physician assessment (81).

With ICS treatment, the time-course for change in symptoms and reliever use is rapid, with statistically significant reductions in less than 1 week (26, 109). Night-waking resolves rapidly, with daytime symptoms and reliever use continuing to improve for up to 7+ months (26) (Figure 1).

**Minimal important difference.** Assessment of minimal important change in diary variables is limited by the choice of comparator. Minimal important differences have rarely been reported for symptom scores, with most papers merely reporting statistically significant changes in mean scores. Santanello and coworkers reported the minimum patient-perceivable improvement in symptom score (scale 0–6, baseline 3) as −0.31, and in reliever use as −0.81 puffs/day (baseline 5.4 puffs/d) (110). For the composite ACQ diary score, a change of 0.5 (scale 0–6, baseline 1.5) was reported as clinically important (111).

**Ambulatory Lung Function**

Ambulatory recording of PEF or FEV₁, or both, provides an objective day-to-day measure of airway obstruction, and is one of the most commonly-reported physiologic outcome variables in clinical trials. Safety has been established, with low rates of maneuver-induced bronchospasm (88, 112). With adequate training, good within-session reproducibility can be achieved for home PEF and FEV₁ monitoring (88, 112), but PEF and spirometric technique should be checked at every visit. Video feedback has been used to improve maneuver quality (113). Subjects should always be asked to complete symptom diaries before measuring PEF, to avoid bias. Some studies exclude the first 1 to 2 days’ data for training effect (114). In children, even with careful training, results from home spirometry are inconsistent (115), so ambulatory lung function monitoring has little role in studies in children.

There are no standardized methods for recording ambulatory spirometric data, but basic guidelines for laboratory-based testing can be applied (116). The exception is that maneuver selection should not be based on the highest FEV₁ + FVC, since FVC from unsupervised maneuvers may not be as reliable as FEV₁ (88). The highest value from three maneuvers is usually analyzed. Most published studies have used morning recordings either on arising or within a specified time-window (e.g., 6:00–8:00 A.M.). However, use of time-windows may increase rather than decrease heterogeneity, as diurnal changes depend more on hours since waking than on the time of day (117). To obtain a more stable measurement, patients are usually asked to record PEF before taking routine study medication, and preferably before reliever medication, but, by contrast with interval clinic spirometry, it is not appropriate to ask patients to withhold as-needed β₂-agonist (118).

In clinical trials, the most commonly-reported diary lung function variable is mean morning PEF. From basic statistical principles, analysis of PEF as L/minute, adjusted within the statistical model for the patient’s age, height, sex, and race, is preferred over analysis as % predicted, so that the relationship between these factors within the study population will be modeled on the actual data rather than fixed by a reference equation derived from other populations. However, for characterization of the study population at baseline, lung function data are most usefully reported as % predicted, with the source of the reference equations stated.

**Responsiveness.** Ambulatory lung function is highly responsive to ICS or LABA treatment, with morning PEF more responsive than evening PEF (118, 119), and morning and evening PEF more responsive than morning and evening FEV₁ (26). With ICS, significant between-group differences can be seen as early as 1 to 4 days (26, 119, 120). With electronic spirometric monitoring, significant between-group differences can be seen even in very mild asthma, where no differences in symptoms are detectable (121). Morning PEF continues to improve for around 2 to 3 months with ICS treatment (122, 123). Inclusion of PEF data recorded after routine (15, 119) or as-needed (118) β₂-agonist use significantly reduces the apparent response to ICS.

**Minimal important difference.** Santanello reported the minimal patient-perceived improvement in PEF as 18.8 L/minute (110) in patients whose baseline FEV₁ was approximately 60% predicted (124), but it cannot be assumed to be the same for patients with better lung function. Most power calculations for efficacy studies quote a clinically relevant difference for morning PEF of 15 or 20 L/minute, but the sample sizes calculated for these studies have ranged from under 40 (125) to almost 200 (126) subjects per group. It is rarely possible to evaluate the appropriateness of published sample sizes, as few papers report the source of the data that were used in the calculations.

**Peak Expiratory Flow Variability**

Variation in lung function represents a different domain of asthma control from static lung function, and both should be reported. There are multiple different calculations of within-day variability (diurnal variability) (127), but the most common is amplitude percent mean (day’s highest minus lowest/mean), averaged over 7 days. With twice-daily PEF measurement, the upper limit of normal for amplitude % mean (95% confidence limit for reference population) is 8% (114) (9.3% in adolescents [127]), increasing to 19% when PEF is recorded more frequently (117). The frequently cited cut-point for diurnal variability of greater than 20% for diagnosing asthma or for classifying asthma as “persistent” (5, 8) is not applicable to twice-daily monitoring. This criterion originated from cosinor modeling of four-times-daily PEF data by patients who had been selected for the presence of definite circadian variability soon after an asthma hospitalization (128).

The most common method of calculation of between-day PEF variability is the lowest PEF (or lowest morning PEF) over 1 or 2 weeks, divided by the highest PEF (8, 69, 123). This index increases as PEF variation decreases.

The above calculations were designed for use in clinical practice, but in clinical trials, PEF variability can also be assessed by standard mathematical methods—for example, standard deviation (51, 129, 130) or coefficient of variation (131). Sophisticated fluctuation analysis has shown associations between PEF variability and risk of exacerbations (131).

**Analysis of Diary Variables**

In the past, treatment effect for diary variables has often been evaluated by ANOVA or paired t tests of the average of the whole treatment period (16) or the last 1 to 2 weeks of treatment (132), with results usually reported as mean or cumulative change from baseline (absolute or percent). However, more recently, mixed model analysis or generalized estimating equations have been used, taking advantage of the power provided by multiple data points per subject (133, 134), and better representing asthma control over the whole treatment period. Use of such statistical methods will reduce the sample sizes required for clinical trials.
Clinical Associations

Although symptoms and lung function are commonly recorded in clinical trials, associations between them are only rarely reported (99). Symptom scores in adults and children generally have moderate or weak correlations with other asthma outcomes, including static lung function, PEF variability, airway reactivity, and airway inflammation (69, 70, 123, 135), consistent with the fact that these represent different domains of asthma control. The lack of correlation may also partly be due to the lack of specificity of asthma symptoms, and to differences in the magnitude and time-course of the response to treatment (26).

Increased PEF variability has been found to be associated with an increased risk of exacerbations (19, 131). Innovative fluctuation analysis of serial lung function measurements is very promising with regard to the prediction of exacerbations, independent of mean PEF (131). This is fueling renewed interest in ambulatory monitoring of PEF and FEV₁ in asthma.

---

**KEY POINTS AND RECOMMENDATIONS:**

**CLINICAL TRIALS**

1. Diaries are useful to assess asthma control in any clinical trial in adults and children, and for patient characterization before randomization, to avoid the problems of patient recall and the effects of change in medication adherence that affect interval questionnaires.

2. Given the evidence for the superiority of data from electronic compared with paper diaries, electronic diaries should be preferred in principle to improve data quality. However, technological reliability and user-friendliness must be ensured. To enhance adherence, participants should be made aware of the recording capability of the device. Given the difference in data quality, electronically collected data should be clearly identified in abstracts and papers.

3. Diaries should be designed at an appropriate reading level, and with features that will optimize adherence by minimizing the burden of monitoring.

4. The use of standardized diaries that have been formally validated in an appropriate population of individuals with asthma (mild or severe, adult or pediatric, self-completed or caregiver-completed) should be promoted.

5. Diaries should include questions on asthma symptoms, night-waking due to asthma, and reliever use. Questions about symptom frequency, intensity, and impact are not interchangeable. The actual wording of diary questions (in the original language) should be provided with all clinical trial reports.

6. “Symptom-free days” is a useful diary variable, but may not be sufficiently responsive in study populations with either very frequent or infrequent symptoms. This variable should be derived from a general question about “asthma symptoms” rather than from several questions about individual symptoms.

7. Diary instructions should advise patients how to record β₂-agonist use. Use of β₂-agonist other than for relief of symptoms (e.g., before exercise) should be recorded and reported separately. Reliever use should be reported both as the proportion of reliever-free days (= β₂-agonist free days) and as the mean of occasions (not puffs) per day, averaged over 1 week.

8. Ambulatory recording of lung function (FEV₁ and/or PEF) is used in some studies to provide information complementary to that provided by symptom diaries or clinic FEV₁. Mean morning PEF provides information about current clinical control, and peak flow variability provides independent information about risk of future exacerbations.

9. In clinical trials, where data are analyzed electronically, standard statistical methods of assessing variability are preferred to previous methods such as amplitude percent mean.

10. Where feasible, diary data should be collected and analyzed over the whole treatment period to capture asthma control over a longer interval. Methods of analysis such as mixed model and generalized estimating equations should be used where possible, to maximize the power of the study and increase the information that is obtained about asthma control between clinic visits.

These points are also applicable for clinical trials in primary care. In addition, in primary care:

- Standardized diaries, suitable for the diverse and often milder patient populations should be validated for clinical trials.
- Symptom-free days may not be an appropriate outcome measure in primary care studies due to lack of sensitivity in mild asthma.

---

**KEY POINTS AND RECOMMENDATIONS:**

**CLINICAL PRACTICE**

1. Symptoms and lung function represent different domains of asthma, and they correlate poorly over time in individual patients, so both need to be monitored by clinicians assessing asthma control in clinical practice.

2. Long-term diaries are not needed for the clinical management of asthma in the majority of patients, but may be relevant in “poor perceivers” (patients who have difficulty sensing airway obstruction) or patients with frequent exacerbations.

3. When patients are carrying out ambulatory lung function monitoring, their monitoring device should also be used for testing in the doctor’s office, to allow comparison with their usual readings.

4. Lung function diary monitoring is to be encouraged in the diagnosis of asthma. The upper limit of normal for amplitude percent mean with twice-daily monitoring is 8%, not the traditionally quoted cut-point of 15 to 20%.

---

**PEDIATRIC ISSUES**

1. Pediatric studies should use diaries specifically validated for this age group.

2. In pediatric studies, as in adult studies, paper diaries are subject to poor adherence and data fabrication.
3. For children less than 12 years of age, diary completion by caregiver rather than child may result in a more complete dataset, but may introduce bias. Use of pictorial symptom diaries may allow self-completion by the child.

4. Ambulatory lung function monitoring has little role in studies in children.

**RESEARCH QUESTIONS**

1. Data are urgently needed on internal consistency and test-retest reproducibility of diary measures, and their correlation with other asthma outcomes, to develop a suite of diary forms, questions, and scores for different purposes (mild or severe asthma, pediatric or adult, parent-completed or child-completed, long-term or short-term use).

2. Minimal important differences need to be defined and quantified, to allow for power calculations if diary variables are used as primary endpoints. Underpinning this work, there needs to be analysis of the extent to which various measures of asthma control predict future risk to the patient.

3. Identification of ways to optimize adherence with diaries, particularly for long-term studies, deserve priority.

4. There is an urgent need for reliable low-cost user-friendly electronic devices and interactive internet-based software, for both symptom and lung function diaries, to eliminate data fabrication and obtain more reliable information about the relationship between diary measures and other variables.

5. Fluctuation analysis of serial measurements of ambulatory PEF or FEV₁ deserves further development and validation in relation to the prediction of disease outcomes.

---

**LUNG FUNCTION AND AIRWAY HYPERRESPONSIVENESS**

**Spirometry**

**Measurement and analysis.** Based on the definition of asthma and the goals of treatment (8, 9), spirometry is one of the fundamental measures of asthma control. It provides an objective and highly reproducible measure of airflow limitation caused by smooth muscle contraction or structural changes. The main spirometric parameters relevant to asthma are FEV₁, FVC (as VC or FEV₆), FEV₁/FVC ratio, bronchodilator (BD) responsiveness (change in FEV₁ after inhaled bronchodilator), and post-BD spirometry.

For clinical trials, pre-BD FEV₁ is defined as FEV₁ recorded after withholding of SABA and LABA for a period appropriate for their duration of action (e.g., 6 h for SABA and 12 h for LABA). Pre-BD FEV₁ has been used as the primary endpoint of lung function in the majority of asthma clinical trials over the last three decades. This had its origins in the early focus on airway obstruction as the primary characteristic of asthma. Pre BD FEV₁ is influenced by short-term fluctuations in airflow limitation and therefore can be considered as a measure of asthma control. Adherence to methods recommended by ATS/ERS (136) minimizes the effect of patient effort. Elderly patients can perform spirometry with good quality (137). Specific ATS/ERS guidelines are available for preschool children (138).

Post-BD FEV₁ is defined as FEV₁ recorded 15 minutes after administration of 400 μg albuterol or equivalent. It is not considered necessary to specify whether LABA or study medication should be withheld, as FEV₁ is close to plateau levels after 400 μg albuterol. Post-BD FEV₁ values are likely to be determined by airway structure and may be used as a measure of severity in describing certain asthma phenotypes. Both pre-BD and post-BD FEV₁ may change with treatment that modifies underlying disease activity. In clinical practice, lung function recordings are usually made without withholding of regular controller medications. The resulting “on-treatment FEV₁” will only be significantly different from “pre-BD FEV₁” for patients taking LABA. To standardize the measurement of “on-treatment” FEV₁, it should be performed after withholding SABA appropriately.

For clinical trials, a more stringent goal for within-test reproducibility of FEV₁, of less than or equal to 100 ml, should be considered than the 150 ml recommended by ATS/ERS for general use (136). The most appropriate spirometry outcome variable is pre-BD FEV₁, adjusted for ATS/ERS for general use (136). The most appropriate spirometry outcome variable is pre-BD FEV₁, adjusted for ATS/ERS guidelines (140) recommend ECSC reference values for Europe (141) and NHANES III reference values for North America (142), including Hispanic values that are also suitable for Central and South America (143). For each lung function index, the lower limit of the normal range is defined by the fifth percentile.

The short-term (< 1 h) within-subject reproducibility for FEV₁ and FVC is very good: less than or equal to 200 ml and 5% for both healthy subjects and patients with asthma (all ages) (144, 145). Values within these limits are indicative of acceptable biological and technical variability. Asthma is characterized by variable airflow limitation, so, as expected, the visit-to-visit reproducibility of FEV₁ (without an intervention) is significantly better in healthy subjects than in asthma (146). However, although intuitively one might anticipate it, there are few data to confirm that between-measure variability in FEV₁ correlates with overall asthma control.

**Responsiveness and time scale.** Based on within-subject reproducibility, an improvement in FEV₁ of greater than or equal to 12% and 200 ml in patients with asthma with baseline airway obstruction is usually considered to be significant (140). In clinical trials, this level of improvement with bronchodilator is...

---

6 h for SABA and
12 h for

often pre-specified by study inclusion criteria, to provide objective confirmation of the diagnosis of asthma (139). FEV\textsubscript{1} is more responsive than FVC, since FVC is usually even closer to its normal or maximal value. Although baseline clinical FEV\textsubscript{1} is often within the normal range in patients with mild or well-controlled asthma (147–149), a ceiling effect for responsiveness of FEV\textsubscript{1} is not necessarily seen in adult populations (121, 150).

The FEV\textsubscript{1} responds to fast-acting inhaled bronchodilators within ten minutes, this effect lasting for at least two hours (151). The FEV\textsubscript{1} also responds to slower onset inhaled bronchodilators (salmeterol and anticholinergics) within about an hour (152). Both pre-BD and post-BD FEV\textsubscript{1} increase within days with ICS (26) or prednisone therapy, and the improvement lasts for days to weeks after discontinuing chronic ICS therapy (153). Measurement of post-BD FEV\textsubscript{1} over many years is recommended for monitoring growth and decline in lung function, because it is less affected by variability in smooth muscle tone. Although post-BD FEV\textsubscript{1} does not give direct information about airway structure, it is the recommended functional measure when airway remodeling is the focus of interest (154, 155).

The minimal important difference (MID) for improvement and worsening in FEV\textsubscript{1}, based on patient perception of change, is about 10% (110, 156), but this is not well established. This difference is greater than the expected test-retest variation (140).

**Associations with other asthma control measures.** Airflow limitation is a major cause of dyspnea and chest tightness, but FEV\textsubscript{1} and symptoms are only weakly associated in asthma. In cross-sectional analysis, for example at study entry, pre-BD FEV\textsubscript{1} (% predicted) has been associated with most other measures of asthma control (3, 61, 157). However, correlations with symptoms are usually weak (158–161), and correlation with disease-specific quality of life is poor (162, 163). Longitudinal changes in FEV\textsubscript{1} have also been associated with changes in most other asthma outcome measures (164–166). However, again, such associations are generally poor, including those with indices of airway inflammation (167).

For assessment of future risk, low pre-BD FEV\textsubscript{1} percent predicted (66, 168, 169) or low on-treatment or random FEV\textsubscript{1} percent predicted (170–172) are strong, independent predictors of subsequent asthma exacerbations. Bronchodilator reversibility is also an independent predictor of death due to asthma (168).

These findings indicate that spirometry provides complementary information that is not provided by other outcome variables (3, 173). Indeed, improvement in symptoms using, for example, LABA therapy, may occur without any change in prebronchodilator FEV\textsubscript{1} (61). Change in FEV\textsubscript{1} is also moderately associated with change in PEF (174). Bronchodilator responsiveness is only weakly associated with measures of AHR and airway inflammation, but is an independent predictor of response to ICS therapy (156, 175). However, most clinical trials have not published these associations (even when they were determined).

**Peak Expiratory Flow**

PEF is inferior to FEV\textsubscript{1} as a clinic-measured parameter of airways obstruction as it confers no advantage in reproducibility, lacks accurate reference values for many populations, and may underestimate airway obstruction in individuals with airway remodeling. Where serial office PEFs are recorded, the same instrument should be used on each occasion if possible. PEF measurement is most suitable for ambulatory monitoring for within-patient comparisons over time, although, with compact electronic devices, daily home monitoring of FEV\textsubscript{1} has become a realistic and successful alternative to this (121) (see **Ambulatory Lung Function**). Interest in ambulatory lung function monitoring has been renewed by innovative fluctuation analysis which has been found to predict clinical course (131) (see [Diary Data in the Assessment of Asthma Control: Peak Expiratory Flow Variability](#)).

**Lung Volumes and Airway Resistance**

Very few asthma clinical trials have used pulmonary function tests other than spirometry or bronchial challenge testing as endpoints. While hyperinflation contributes to sensations of dyspnea and chest tightness during asthma exacerbations or exercise (176), there is no convincing evidence to date that measuring reductions in RV or FRC adds clinically important information to the increases in FEV\textsubscript{1} or FVC that occur simultaneously with successful therapy.

Airway resistance is increased during asthma exacerbations and falls with successful asthma therapy. It can be measured with body plethysmography or other instruments (forced oscillation or interrupter technique). Advantages of airway resistance measurements over spirometry include: (1) no need for breath-holding maneuvers are not needed, (2) the effects of deep inspiration are eliminated (177), and (3) results can be obtained from preschool aged children (178, 179). Methods for the forced oscillation technique (FOT) have been standardized by the ERS (180). However, airway resistance tests have some disadvantages, including equipment costs, few validation studies, the need for trained technologists, and the lower signal–noise ratio compared with FEV\textsubscript{1} measurements (181).

**Airway Hyperresponsiveness**

**Method and parameters.** Airway hyperresponsiveness (AHR) is an objective, well-standardized measure of variable airflow limitation. It reflects the increased sensitivity of the airways to inhaled stimuli, a problem reported by the majority of patients with asthma, even when spirometry is normal. It can be measured by “direct” and “indirect” challenge tests (182, 183), which refer to the mode of action of the agents in relation to smooth muscle contraction. Methacholine chloride and histamine diphosphate are most commonly used as direct smooth muscle stimuli, but the results cannot be used interchangeably (182, 184). Currently, the most frequently used indirect stimuli, which involve multiple cellular pathways, are hypertonic saline, adenosine monophosphate (AMP), and mannitol (182, 183, 185).

Guidelines for challenge testing have been published by the ERS (182) and ATS (184). There are two safe and validated methods for inhaling aqueous solutions of pharmacologic stimuli. These are the 2-minute tidal breathing method (182–184), and the dosimeter method (182–184). These methods are both well standardized, but they cannot be used interchangeably (186). New commercially available dosimeter methods are emerging, and may offer an alternative provided that the methodology is validated against the two gold standards above.

Regardless of the method, only the first part of the sigmoid log concentration or log dose–response curve can be recorded. The curve is usually expressed by its position: the provocative concentration/dose (PC or PD) to cause a certain degree of airway narrowing (e.g., 20% fall in FEV\textsubscript{1}: PC\textsubscript{20}, PD\textsubscript{20}) as measured by log-linear interpolation (182, 183). It is customary to use PC\textsubscript{20} for the tidal breathing method, and PD\textsubscript{20} for the dosimeter method (although the dose delivered to the Airways can only be roughly estimated). A 20% fall in FEV\textsubscript{1} can often not be obtained in control subjects without asthma, for which the so-called “two-point slope” method is a valid alternative.
(187). Extrapolation of dose–response curves is strongly discouraged. In young children, measurement of respiratory resistance may be an alternative to FEV1 during challenge testing. This needs further validation (138, 188).

Challenges are safe when the protocols comply with current recommendations (182–184), and are well tolerated even in young children (138, 189). Detailed standardized laboratory operating procedures should be at hand. Compulsory safety precautions (182–184) must be obeyed at all times.

**Reference values and reproducibility.** The normal range of PC20 for both methacholine chloride and histamine diphosphate is greater than or equal to 8 mg/ml (190), while for PD20 it is greater than or equal to 7.8 µmol for both agents (191). The normal range of two-point slope to methacholine has been reported to be less than 2.39% fall/µmol, but appears to be a function of age and pre-test FEV1 (192). The normal range of PC20 for AMP has been suggested to be greater than or equal to 200 mg/ml (193). The PC20 and PD20 are adequately described by their geometric mean values and geometric standard deviations (expressed in doubling concentration or dose, respectively).

The 95% confidence interval for short-term repeated measurements of PC20 and PD20 for histamine and methacholine is ±1.5 doubling dose (182), and for AMP ±1.7 doubling dose (194). The intraclass correlation coefficient for repeated measurement of the two-point slope to histamine is only 0.26, but can be improved by calculating the slope by least-square analysis (195). Recommendations for sample size estimations based on PC20 and PD20 are available (196).

**Responsiveness and time-scale.** Bronchial challenge tests are immediately responsive to pretreatment with functional antagonists, such as short-acting (197) and long-acting (198) β2-agonists. Short-acting bronchodilators should be stopped 8 hours and LABA 36 hours before the test (182–184). Even after withholding of salmeterol for 24 hours, there is still a greater than twofold shift in PC20 (198). In clinical trials, the impact of these drug-free intervals on the measurement of other outcome variables needs to be considered.

Challenge tests are highly suited for monitoring therapy aimed at disease modification, such as antiinflammatory therapy. Increases in PC20 and PD20 with ICS are dose- and time-dependent. Meta-analyses show on average 0.9 doubling dose improvement in short-term studies (2–8 wk) (199), with greater improvements by high-dose (>1,000 µg/d beclomethasone equivalent) versus low/medium doses ICS (<1,000 µg/d) (200). Longer term studies have demonstrated much larger average improvements in PD20, reaching 4 doubling doses (16.5-fold change) after 18 months of treatment (26).

The time-scale of changes in PC20 and PD20 in response to ICS therapy varies with the challenge agent. The improvement in PD20 to histamine in adults and to methacholine in children did not show a plateau after 18 months and 22 months of therapy, respectively (26, 201). Inhaled corticosteroids have a much more rapid onset of action when measured by direct challenges such as AMP as compared with methacholine challenge (202, 203). This illustrates that indirect challenges are somewhat more closely dependent on the acute state of inflammatory pathways in the airways (203), whereas methacholine responsiveness additionally appears to be determined by airway dynamics and structure (204). This distinction can be important when monitoring long-term therapy.

**Clinically relevant differences.** During seasonal allergen exposure (205) or after respiratory virus infection (206), group averages of change in PC20 or PD20 of at least 1 doubling dose have been observed. As indicated above, PC20 or PD20 respond to ICS on average by at least 0.9 doubling dose after short-term treatment (199), while responses less than 1 doubling dose were labeled as “poor” in a multicenter study aimed to compare responsiveness of clinical markers to ICS in asthma (156). These data provide an indicative rather than validated estimate of clinically relevant differences.

**Association with other markers and disease outcome.** Airway hyperresponsiveness has a high negative predictive value for the diagnosis of asthma (207, 208). However, in cross-sectional studies, AHR is only weakly associated with symptoms, lung function, and markers of airway inflammation (3). This may be its strength, because it provides independent and complementary information (209). Therefore, AHR is neither a surrogate for clinical symptoms nor for airways inflammation.

Longitudinal studies, mostly performed with direct challenge agents, have demonstrated that hyperresponsiveness is strongly related to the clinical course of asthma. In young infants it can predict the development of asthma later in life (210). As a predictor of future risk, increased hyperresponsiveness predicts loss of control in children (211) and adults (212) with asthma, and appears to be a significant and independent risk factor for the development of irreversible loss of lung function (213–215). Even asymptomatic AHR in the general population appears to be a significant risk factor for subsequent development of wheeze, physician-diagnosed asthma, chronic cough, chronic bronchitis, and COPD (216).

With regard to treatment monitoring, baseline AHR to methacholine predicts the spirometric treatment response to ICS of individuals with asthma (217). Interestingly, individualized treatment additionally guided by AHR to methacholine leads to fewer uncontrolled episodes in adults with asthma (39). This suggests that AHR is an indirect but meaningful marker of asthma control. In the same way, AHR-guided therapy can prevent a decline in lung function in children with asthma (218). Novel interventions specifically targeting bronchial hyperresponsiveness need further exploration (219, 220).

### KEY POINTS AND RECOMMENDATIONS: CLINICAL TRIALS

**Spirometry**

1. Spirometry, as measured by pre-bronchodilator FEV1, is one of the fundamental objective measures of asthma control.
2. Pre-bronchodilator FEV1 is a strong independent predictor of future risk of exacerbations.
3. Spirometry (pre-bronchodilator and post-bronchodilator) should be measured at the baseline examination of most asthma clinical trials, for all study participants aged 6 years or more.
4. The relative importance of spirometry among the endpoints in clinical trials depends on the study objectives. Spirometry provides information on asthma control and future risk complementary to that obtained from symptoms and biomarkers (Tables 1 and 2).
5. FEV1 should be included as a primary endpoint for studies of bronchodilator therapy. FEV1 is highly responsive to the successful relief of bronchoconstriction over the entire range of asthma severity, except for those with normal baseline lung function.
6. Post-bronchodilator FEV\textsubscript{1} is recommended in studies of long-term decline in lung function and airway wall remodeling.

7. The 2005 ATS/ERS guidelines on performance of spirometry should be followed. Automated spirometers with test quality checks, and centralized quality assurance programs, are recommended for use in clinical trials.

8. FEV\textsubscript{1} may be used to characterize the study population. This will also facilitate comparisons with previous studies.

**Airway Hyperresponsiveness**

1. AHR should be regarded as an integrative disease marker, reflecting multiple pathophysiologic mechanisms.

2. The 1993 ERS and 1999 ATS standardized methods should be used. The necessary safety precautions must be followed at all times.

3. Preference should be given to challenge agents produced under Good Manufacturing Practice (GMP) conditions.

4. Where possible, AHR should be included in clinical trials at baseline to characterize the study population.

5. AHR is a desirable outcome in studies focusing on modification of underlying disease activity. Direct challenge agents can be considered for assessing mid- and long-term disease modification, while indirect agents are relatively more responsive when investigating short-term responses to antiinflammatory interventions.

6. AHR can be used as a predictor of future risk of exacerbations and decline in lung function in longitudinal studies of childhood and adult asthma.

**For Clinical Trials in Primary Care**

1. The high standards for quality control and reporting for spirometry that are recommended in this review should also apply to studies in primary care.

2. FEV\textsubscript{1} is preferable to PEF as a clinic-measured physiologic parameter in asthma clinical trials, although PEF may have a place in ambulatory monitoring.

3. The current safety recommendations for bronchial provocation testing preclude its use in most primary care settings.

**KEY POINTS AND RECOMMENDATIONS:**

**CLINICAL PRACTICE**

1. Objective lung function measures should be accurately performed and recorded in the diagnosis and assessment of asthma in primary care, but normal lung function does not exclude a diagnosis of asthma.

2. Lung function measurements in primary care should be performed on appropriate equipment by trained personnel, with monitoring of quality control. Actual values should be entered into the medical record (not just “normal” or “abnormal”).

3. Peak flow measurement in primary care for within-patient comparisons should be performed with the same meter on each occasion.

4. As outlined in the section on diary measures, peak flow variability may assist in confirming the diagnosis of asthma. The upper limit of normal for amplitude percent mean with twice-daily monitoring is 8%, not the traditionally quoted cut-point of 15–20%.

5. AHR may be used to guide asthma therapy given the benefits in reducing exacerbations and decline in lung function.

**PEDIATRIC ISSUES**

1. In children, measuring spirometry is important not only for assessing asthma control, including acute exacerbations, but also for assessing lung development over time.

2. Spirometry can be routinely measured in children aged approximately 6 years and older. However, with appropriate training, preschool children may be able to perform spirometry.

3. Young children have difficulty performing the 6-second forced expiratory maneuver recommended for spirometry in adults. Therefore, shorter expiratory times may be acceptable if reproducible.

4. Forced oscillation procedures and interrupter resistance (R\textsubscript{int}) to measure airways resistance can be applied in children as young as 3 years of age.

5. Challenge tests for AHR in children with asthma require satisfactory cooperation with spirometry. Many children are unwilling to undergo the repeat spirometry required for challenge testing.

**RESEARCH QUESTIONS**

**Spirometry.**

1. The associations between change in lung function and change in other asthma outcomes (especially dyspnea) should be explored using the large databases accumulated from multicenter clinical trials.

2. The relationships between changes in airway structure and measures of airway function require further investigation.

3. The relative utility of electronic PEF monitoring and spirometry in the diagnosis and monitoring of asthma in the community needs to be evaluated more fully.

**Airway Hyperresponsiveness.**

1. Standardization—the ERS and ATS guidelines for bronchial challenge tests require updating, to include GMP requirements, novel challenge agents, and new aerosol dosing techniques.

2. Methodology—quicker and simpler methods should be validated. Is mannitol testing suitable for measuring AHR in primary care settings?

3. Mechanisms—does AHR assess the potential for airway narrowing, airway relaxation, or both?
4. Utility—is an integrative marker such as AHR preferable to single cell or single molecule biomarkers when monitoring disease control? Do the mechanisms of direct and indirect AHR vary between different asthma phenotypes? Can this be used for selecting and assessing therapy?

5. Predictive value—is maximal, long-term improvement of AHR required for optimal asthma control, prevention of exacerbations, and reduction of the decline in lung function? Do direct and indirect challenges have similar predictive value?

6. Is elimination of AHR by novel treatment modalities the way to cure asthma?

**COMPOSITE SCORES FOR ASSESSMENT OF ASTHMA CONTROL**

The concept of composite measures for interval assessments of asthma control is based on three facts: the generally poor correlation between different domains of asthma, the absence of a single “gold standard” for the measurement of asthma control, and evidence that a composite comprising different endpoints provides a more complete picture of asthma control than any single endpoint (158, 221–226). For example, in a retrospective analysis, Bateman and colleagues showed that patients who achieved well-controlled asthma according to a guideline-based composite measure achieved greater improvements in quality of life than if success in only a single component of asthma control was achieved (227).

As in other conditions, the selection of endpoints must be underpinned by knowledge of the clinical manifestations, pathophysiology and natural history of asthma, and of the changes that occur with treatment. Although there is not complete agreement on which parameters should be included in a composite measure for asthma, there is increasing evidence supporting their use to evaluate the efficacy of pharmacologic and nonpharmacologic interventions as well as in clinical practice (52, 81, 111, 157, 163, 221, 227–229). Evidence relating to their use in monitoring asthma control in individual patients is much more limited. This has the potential to be problematic because, although it is assumed on the basis of group results (26, 230, 231) that each component will trend in the same direction in response to improving or deteriorating asthma control, this is not necessarily the case in individual patients or with all treatments (61). In practice, the clinician might intuitively weight an individual component measure in making a therapeutic decision. However, weighting has not, in general, been tested in the use of composite scores in the clinical trial setting. A situation in which change in a composite score is driven predominantly by change in only one of its components may have little clinical validity, either in terms of treatment effect or in terms of importance to the patient. Thus, interpretation of composite scores requires that data for their individual components should also be reported, and reliance should not be placed solely on the composite measure (232).

**Categorical versus Continuous Measures of Asthma Control**

Terms such as “poorly controlled” and “well-controlled” have been used to describe the overall status of individuals or groups of patients. However, this leaves a gap or “no man’s land,” where patients are neither poorly controlled nor well controlled. Individual variables, too, may be couched in positive or negative terms (e.g., “asthma control days” or “asthma exacerbation days”). However, loss of control is not the simple converse of adequate control. Two treatments may be equivalent in terms of well-controlled weeks, but differ significantly with respect to poorly controlled weeks (16, 33). Some of these problems can be avoided by expressing control as a continuous numeric variable, as described below. This approach simplifies recording and makes self-evaluation by patients possible.

The process of creating such composite scores for asthma has been problematic because it has been largely empiric. It has involved transforming several originally continuous variables that had clinical validity (e.g., symptoms, reliever use) into individual categories to which scores were assigned, then summing these scores to create the composite score. Often, categorical cut-points were then applied to the total score, to allow categorization of patients into well-controlled, poorly controlled, and so on. This may facilitate validation of the composite score against categorical measures such as physician assessment. However, even when results are expressed as a continuous numeric variable, there is again an indeterminate zone between values that provide a high level of certainty that a patient is not well controlled and those that are considered to represent adequate control (233) (see Table E1 in the online supplement). Ideally, the appropriateness of individual cut-point values for each component should be validated, both alone and as part of the composite measure to which it contributes (81, 225).

**Group versus Individual Data**

Care is required when interpreting the results of composite measures that are presented as group data. For example, the proportion of “asthma control days” for a study population may increase from 21 to 50% (109), but this provides no indication of the adequacy of control in individuals (221). It is also susceptible to the “ceiling effect” if a large proportion of patients achieve good control. The proportion of patients that achieve the target level of control provides a more satisfactory indication of the success of treatment (52, 221).

**Time as a Factor in Composite Measures of Control**

Like other outcome measures, composite scores which are derived from daily variables must be indexed to a clinically meaningful period of time. Existing composite tools use assessment periods varying from 1 week to 1 month (158, 223, 225, 226) (Table E1). To provide a summary of the interval between infrequent clinic visits, composite scores may be recorded in a diary. The rate of change also needs to be considered: different symptoms may not disappear at the same rate. For example, night waking disappears first as control is achieved, and is not the first to return when control is lost. This difference in responsiveness earns night waking a separate place in most composite measures of control (231).

**Composite Measures Expressed as Categorical Variables**

Composite measures expressed as categorical variables include asthma control days, asthma-free days, episode-free days, and exacerbation-free days. They share certain features: empiric derivation, lack of standardization, and the limited information they provide on control in individual patients. However, they have been widely used in asthma clinical trials. Asthma control days/asthma-free days/episode-free days. The simplest composite measures are asthma-free days or asthma-control days, typically defined as days with no symptoms, no night-waking, no reliever use, and no exacerbation (221, 222, 234). However, different studies have allowed (235) or excluded (236) days with pre-exercise \( \beta_2 \)-agonist use. Some have also excluded asthma “attacks” and need for additional therapy (24, 108, 234, 237). Some authors have permitted some use of
bronchodilator as reliever without that day losing its asthma-control status (237). This approach is potentially problematic: it is arguable that an asthma-free day ought to be one in which reliever use is zero, even though guidelines do not insist on such a stringent criterion for well-controlled asthma.

**“Well-controlled” asthma weeks and “Total Control” weeks.** In many studies, the method for scoring asthma control days has been lengthened to periods of a week (33, 52, 221) or several weeks (52, 227), referred to as “well-controlled weeks” or as a period of “well-controlled” asthma or “total control” (52, 227). Threshold values for each of these outcome variables, which are usually derived from goals of treatment in asthma guidelines (8, 9, 238), are mostly arbitrary. Patients are usually required to achieve control in each outcome and/or for all the days of the week, for most or all the weeks of the assessment period (52, 227).

**Guideline-based categories of asthma control.** Several recent national and international guidelines have incorporated classification systems for asthma control, based on clinical consensus. These have used many different terms; for example, the GINA Report of 2006 used categories of “Controlled,” “Partly Controlled,” and “Uncontrolled” (8); the NHLBI Expert Panel Report 3 used “Well Controlled,” “Not Well Controlled,” and “Very Poorly Controlled” (9); the French guidelines used “Optimal,” “Acceptable,” and “Unacceptable” (228); and a New Zealand report used “Optimal Control,” “Sub-Optimal Control,” “Not Well Controlled,” and “Markedly Out of Control” (239). Different criteria and cut-points have also been used in each of these classification systems.

**Composite Measures Expressed as Numeric Variables**

The concept of representing control as a numeric score is attractive and has been adopted by several research groups. Several measures have been derived. The obvious advantage of numeric composite measures is that absolute values as well as changes in numeric scores are relatively easy for patients and carers to understand and record. With time, patients’ awareness of the difference between satisfactory versus unsatisfactory control can be related to the numeric scores. They may be also more user-friendly for use in programs of self-management (240).

Examples of numeric composite measures are the ACQ (158), the Asthma Control Test (ACT) (225), the Asthma Treatment Assessment Questionnaire (ATAQ) (226), and the Asthma Control Scoring System (ACSS) (223) (Table E1). Weighting of items has not been examined for any of these composite scores. Comparisons between such composite scores are difficult because each uses a different “gold standard” for what constitutes good asthma control, and they have been directly compared in only a few studies (241–243).

Before using composite scores, it is important to check copyright restrictions and charges.

**Asthma Control Questionnaire.** The ACQ (158) was developed by Juniper and coworkers for assessing asthma control in clinical trials and clinical practice. Questions based on recall of the previous 7 days comprise breathlessness, nocturnal waking, symptoms on waking, activity limitation, wheeze, frequency of SABA use, and pre-bronchodilator FEV1% predicted. All seven items are scored on a 7-point scale without weighting (0 = good control, 6 = poor control) and the overall score (range, 0–6) is the mean of the responses. The ACQ has been validated against quality of life and physician global assessment (81, 111, 158, 163, 233), and the MID is 0.5 (111). The optimal cut-point for “Well-Controlled” using the Gaining Optimal Asthma Control (GOAL) classification is less than or equal to 0.75, and a value of greater than or equal to 1.50 confirms “not Well-Controlled” asthma (233). Shortened versions, with omission of SABA use and/or FEV1, perform almost as well as the 7-item version (111, 233) and may be suitable for self-completion in primary care (224), or in patients taking LABA. Wording of the validated ACQ is slightly different from the originally published version (158).

A shortcoming of the ACQ is that most patients’ scores are less than or equal to 2.5, with scores of greater than or equal to 4 only occurring with severe exacerbations. This suggests that the range and intervals for individual item scores could be improved. Also, the response scales may be more complex and time-consuming than is necessary, and its acceptance for use in primary care needs to be demonstrated. Although ACQ includes pre-bronchodilator FEV1, a predictor of risk of exacerbations, change in this component may be outweighed by the remaining six symptom/reliever components, as was seen in one study of LABA monotherapy (61).

**Asthma Control Test.** The ACT was developed by Nathan and colleagues (225) and is a trademark of QualityMetric (Lincoln, RI). This self-completed instrument comprises five items: shortness of breath, night-time waking, interference with activity, rescue bronchodilator use, and patient rating of asthma control, over the past the month. Each item is scored using a 1–5 scale and then summed (total score, 5–25). The test is easy to use, and can be easily completed on the internet (244) or telephone (245). A pediatric version (C-ACT; range, 0–27) has been developed for children aged 4 to 11 years (246). It comprises four questions for the child (how is your asthma today?; exercise-induced symptoms; cough; and night waking) with picture prompts for responses, plus three questions for the parents (days in the last month with daytime asthma symptoms; wheezing; and night-time waking because of asthma).

The ACT has been validated against specialist’s rating of control and spirometry (225, 240, 243), and quality of life (247). The published cut-points for well-controlled asthma and poorly controlled asthma are greater than or equal to 20 and less than or equal to 15, respectively (225, 243, 247); for the C-ACT, a score of less than 20 corresponds to uncontrolled asthma (246). A change in ACT score of 1.88 corresponds to a change of one level in physician rating of asthma control (243). Although the ACT has been translated into more than 40 languages, further research is required to ensure validity in different languages and practice settings. The length, presentation, and ease of use of the ACT make it an attractive option as a self-assessment tool both in its paper and web-active versions. Its utility in research requires further assessment.

In a comparative study, ACT and ACQ showed similar reproducibility, discriminant validity, and sensitivity/specificity for detecting poorly controlled asthma (243).

**Asthma Therapy Assessment Questionnaire.** The ATAQ is a self-administered questionnaire that was developed by Vollmer and coworkers (226) for assessing asthma control in patient populations for health utilization and planning purposes (226, 248, 249) and for clinical use in health maintenance organizations (250). The ATAQ includes a four-item control questionnaire (itself often being referred to as ATAQ) and a longer questionnaire about barriers to asthma management. The control “domains,” relating to the previous 4 weeks, are: self-perception of asthma control; missed work, school, or normal daily activities due to asthma; night-time waking due to asthma symptoms; and excessive use of “quick relief” inhaled medication. Answers are dichotomized, leading to a composite score ranging from 0–4, with 0 representing no control problems and 4 indicating 4 control problems. Each question has the same weighting. An ATAQ for children and adolescents (range, 0–7) has been developed (251).
The ATAQ has been validated against major quality-of-life instruments (Medical Outcomes Short-Form 36 [SF-36], St. George’s Respiratory Questionnaire [SGRQ], and AQLQ) and has been shown to predict need for acute care for asthma (248, 252). The questionnaire is very simple to use. Reports of its use in clinical practice and clinical trials are awaited.

**Asthma Control Scoring System.** Boulet and colleagues developed the ACSS with the intention that it should be a simple, more “flexible” approach to documenting asthma control (223). The component parts are: clinical (daytime and nocturnal symptoms, rescue SABA use, and activities, each during the last 7 d), scored by the patient; physiologic (FEV₁); and induced sputum eosinophil count. It is the only composite instrument to include a marker of airway inflammation. The overall score is the average of the available components expressed as a percentage, with 100% representing ideal control. However, the authors proposed that “the respective weight of each component could be interpreted differently by clinicians,” opening the way for different formulae for arriving at a global control score (253).

A small study examined the ACSS against the Mini-AQLQ and ACQ in 44 patients with asthma (223), and its measurement properties were reported from another small study (n = 28, of whom 28 had sputum data) (253). It remains to be established in a larger study whether the systematic or occasional exclusion of the inflammatory component of the ACSS alters its measurement properties and clinical utility.

**Summary**

Composite measures are designed to capture different and often independent aspects of asthma control. They attempt to provide a summary statement about a complex clinical state. Most of the current scores focus on current clinical control rather than the underlying disease activity. Their use in clinical trials enables numeric comparisons of treatment effects to be made. However, their interpretation in clinical practice is not necessarily straightforward, particularly where the response of individual components may be discordant or some items may be missing. Considerable progress has been made in validating these instruments for research and clinical use. Whether any improvements would be achieved by weighting of their individual components has not been explored, particularly with regard to predicting future risk.

**KEY POINTS AND RECOMMENDATIONS:**

**CLINICAL TRIALS**

1. Composite scores have the potential to be used as primary or coprimary endpoints in clinical trials.
2. Composite scores should be relatively simple and easy to administer. They must be suitable for a full range of patients, and modified to suit different patient groups (e.g., children).
3. Composite measures that record asthma control as a simple numeric form should be favored. The number of individual items in composite measures should be kept to a minimum, and the contribution of each item to the performance of the instrument must be evaluated.
4. Existing composite measures have primarily been validated against physician judgment or other measures of current clinical control without weighting of their components. Against this standard, physiologic data such as spirometry appear to contribute little in respect of either responsiveness or precision.
5. Measures that describe achievement of control both in terms of completeness and duration (e.g., well-controlled weeks) and measures of departure from control (e.g., poorly controlled weeks or exacerbations) provide complementary but differing information and may be used together. Results for these outcome measures should be expressed both as group means and also as the proportion of participants who achieve the target level of control.
6. Categorical composite measures do not necessarily provide a full picture of asthma control, particularly with regard to future risk.
7. Recommendations about composite measures are limited by the relative lack of validation in a wider range of settings, over longer periods of follow-up and in patients with different asthma phenotypes, different levels of control, and on different types of treatment.
8. Investigators should check conditions of use and charges before using any of the existing composite scores.

The above statements also apply in clinical trials based in primary care.

**KEY POINTS AND RECOMMENDATIONS:**

**CLINICAL PRACTICE**

A composite measure does not provide full information on a patient’s current clinical state. Clinical consultations should continue to include additional questions that address other aspects of the impact of asthma on the individual.

**PEDIATRIC ISSUES**

1. Assessment of asthma control in children is usually based on parent reports. The reliability of symptom assessment from questionnaires may be influenced by poor symptom perception and reporting by the child or by the parents.
2. Symptom-free days is a useful endpoint for pediatric asthma studies, and easier to record than symptom scores.
3. Since the decision to take rescue bronchodilator is often made by the parents, this indicator may not assess asthma control accurately in young children.
4. Some age-specific pediatric versions of commonly used control questionnaires, validated in different languages, are becoming available for clinical studies (e.g., ATAQ, ACT). However, their potential application in pediatric clinical practice needs to be carefully evaluated.
5. In children, the target level of control aimed for is influenced by safety concerns about the long-term use of high doses of ICS in this age group. In this context, a careful assessment of risk and benefits may lead to acceptance of less than complete symptom control of asthma.

**RESEARCH QUESTIONS**

1. The measurement properties of composite measures should be validated both in clinical trials (groups of
patients) and in large prospective studies in “real life” settings (individuals) including in primary care, to ensure that they provide content validity as well as reflect clinically meaningful outcomes.

2. Further studies are required to establish whether prospective monitoring of asthma control using composite measures improves asthma outcomes and/or predicts future risk, and whether adding physiologic or inflammatory components to composite scores adds value in particular phenotypes, for example in poor perceivers or patients with severe asthma.

3. There is a need for further studies to investigate the most useful composite measures for primary care settings.

**Biomarkers of Airway Inflammation**

Over the last 15 years there has been increasing interest in the noninvasive assessment of airway inflammation as an adjunct to the assessment of clinical asthma control (254–256). A number of candidate measures have been developed and validated (254, 257–259). Some have been evaluated in clinical trials, and there is increasing evidence that the information provided by noninvasive markers results in more effective use of available asthma treatments (260–262).

The assessment of airway inflammation is moving rapidly, but not all new techniques have been developed to the point where their clinical utility has been validated. We have selectively focused on the methodological aspects of induced sputum analysis, measurement of the fraction of exhaled nitric oxide ($\text{FE}_{\text{NO}}$), exhaled breath condensate (EBC) analysis, and the use of serum eosinophilic cationic protein (ECP). Indirect assessment of airway inflammation using the peripheral blood eosinophil count has a long pedigree and showed promise in earlier studies (263, 264). However, there has been little recent work on this marker and it will not be considered further.

**Induced Sputum**

**Measurement methods and interpretation.** The methodology for sputum induction and processing was reviewed by an ERS Task Force in 2002 (264). Induced sputum is not possible in children aged less than 8 years. Most centers report success rates, defined as obtaining a readable sputum cytospin, of 80 to 90% in adults, and somewhat lower in children (265). Low baseline lung function ($\text{FEV}_1 < 1.0$ $\text{L}$) is a relative contraindication, and in all patients, pretreatment with SABA is recommended to avoid inducing bronchospasm. Occasionally, patients experience excessive coughing culminating in vomiting during sputum induction.

The protocols for sputum induction differ mainly in the output of the ultrasonic nebulizer used, and sputum processing (sputum plugs or whole sample selection). Usually sputum is induced with nebulized hypertonic saline and processed with the aid of the mucolytic dithiothreitol. There is no evidence that different methods result in clinically important differences in success rates or sputum differential cell counts (264, 266).

The principal readout from induced sputum is the differential inflammatory cell count, expressed as a percentage, based on a manual count of 400 inflammatory cells (eosinophils, neutrophils, macrophages, lymphocytes, and epithelial cells) on a stained cytospin preparation. The total cell count, cell viability and squamous cell contamination should also be reported (264). Inflammatory cell counts can also be expressed as total count (i.e., total cell count × proportion of that inflammatory cell). There is an approximately linear relationship between the differential and total count up to a differential count of 80%. Above this level, total counts may provide more information on the intensity of the inflammatory response; however, they are less repeatable than differential counts (254, 255). This may be particularly important when assessing neutrophilic inflammation.

The induced sputum supernatant can be used to assay molecular markers of inflammation. In general, this technique is more successful for effector mediators than cytokines, and for Th-1–associated cytokines compared with those associated with Th-2 responses (264). Assay of sputum supernatant mediators has made an important contribution to our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of the mechanisms of airway disease, but there is no evidence that these measures inform our understanding of.
Fractional Concentration of Exhaled Nitric Oxide

The exact relationship between the fractional concentration of exhaled nitric oxide (\(F_{\text{ENO}}\)) and the underlying pathologic process in asthma remains unclear (292). The increasing use of \(F_{\text{ENO}}\) as a surrogate marker for the presence of clinically relevant eosinophilia is based on significant correlations between \(F_{\text{ENO}}\) measurements and eosinophilic airway inflammation (293), although the scope for both false positives and false negatives remains significant (294).

**Measurement methods and interpretation.** This subject has been reviewed by several Task Forces (257, 295, 296). There is some concern that absolute values may differ between different analyzers, which has obvious implications for normal ranges (315, 316), and stability over time needs to be established. The between-center variability in \(F_{\text{ENO}}\) is likely to be similar to the within-subject variation if the same type of analyzer is used.

**Responsiveness.** \(F_{\text{ENO}}\) is increased (by about 60%) during the late response to allergen in subjects with atopic asthma (317), and is reduced two- to fourfold by corticosteroids in patients with asthma (279, 318, 319). There is conflicting information on the dose–response relationship between ICS dose and reduction in \(F_{\text{ENO}}\). Available data are consistent with a dose-related effect with low dose treatment but no additional benefit at a group level above a budesonide dose of 400 \(\mu\)g/day (279, 320, 321). The time scale of response to ICS has not been adequately addressed; estimates for the time to a measurable effect range from 3 days (321) to 8 weeks (318). Based on the available data, an estimate of the clinically meaningful change would be twofold in either direction.

**Associations with other outcome variables.** \(F_{\text{ENO}}\) is a reasonably robust estimate of the presence of eosinophilic airway inflammation across a wide range of patients differing in diagnosis (294). The association is lost in current smokers (294). There is accumulating evidence of an important degree of dissociation between eosinophilic airway inflammation and symptoms/disordered airway function in some asthma phenotypes (3, 4, 260, 288).

Corticosteroid reduction studies have consistently shown that a raised sputum eosinophil count is predictive of the development of an exacerbation (212, 277, 282, 289), and management strategies aimed at normalizing sputum eosinophil counts have been associated with up to 60% reduction in severe asthma exacerbations (260, 261, 290), with particular benefit in patients taking LABA or those with more severe asthma (261). In this population, induced sputum analysis has been shown to be cost effective (260). Although the majority of patients in primary care demonstrate concordance between symptoms and eosinophilic airway inflammation, significant discordance is seen in patients referred for secondary care (288). Cluster analysis has demonstrated that with sputum-guided therapy, the majority of benefit in reducing exacerbations occurs in patients with inflammation-predominant asthma, whereas the majority of benefit in reducing ICs dose is seen in patients with predominant symptoms and little inflammation (288). The differential neutrophil count has been shown to relate inversely to the post-bronchodilator FEV\(_1\) (293), although the scope for both false positives and false negatives is based on significant correlations between \(F_{\text{ENO}}\) measurements and eosinophilic airway inflammation (293), although the scope for both false positives and false negatives remains significant (294).
pediatric studies, FeNO predicted the need for continuing ICS therapy in patients whose asthma appeared stable: after ICS withdrawal, increasing FeNO predicted loss of control (322, 323), whereas persistently low FeNO predicted successful ICS withdrawal (323). Michils and coworkers reported that in patients with mild asthma, a decrease in FeNO of greater than 40% had a positive predictive value for improved ACQ of 83% (324). Five studies have investigated asthma outcomes when ICS dose was guided by FeNO (262, 325–328). In one study, this approach resulted in lower ICS doses compared with standard clinical management (262). A study in children showed that a management strategy based on FeNO was associated with improved AHR (325). All three studies reported fewer exacerbations when treatment was guided by FeNO, but the results were not statistically significant (262, 325, 326). Individually, the outcomes of these studies may be due to problems with study design or power, but overall the conclusion is that optimizing ICS dose is not a primary indication for using FeNO in uncomplicated asthma.

Exhaled Breath Condensate

Methods and interpretation. The collection of EBC and subsequent analysis of inflammatory markers is a more recent development in noninvasive asthma monitoring technologies. Cooling of expired air condenses exhaled breath, which contains water vapor, respiratory droplets, and particles. This is collected and assessed using conventional assays. Methodologic aspects of EBC collection and analysis have been reviewed by an ERS working party (258). Current methods for the collection of EBC vary primarily in the type of condenser. The physical surface properties of each condenser system may influence the condensate that is collected, and it is possible that there may be great variation in particles collected by each system. The influence of salivary contamination on EBC values may be considerable, as many of the mediators assayed are found in high concentration in saliva (329).

Many markers of airway inflammation have been reported in EBC. The most commonly reported markers include those that indicate oxidative stress such as 8-isoprostane and hydrogen peroxide, as well as the leukotrienes (cysteinyl and B4) and airway pH. Other less frequently reported markers include cytokines such as IL-6. The measurement characteristics of the common markers are shown in Table E3. The application of non–hypothesis-driven “metabolomics” may assist in clarifying the relationship between EBC and other markers of asthma control (330).

Serum Eosinophil Cationic Protein

Methods and interpretation. A standardized collection, processing, and testing method has been described for ECP (331). Critical factors include storage temperature and the time to analysis. Serum ECP concentrations are higher than ethylenediaminetetraacetic acid plasma concentrations, probably because blood eosinophils continue to produce ECP ex vivo in the absence of additives. Serum ECP concentrations are preferred, as they appear to be better at discriminating health from disease (331). Circadian variation of serum ECP concentrations is present, indicating the need to standardize collection time (332), and there is evidence that a promoter polymorphism is a major determinant of serum ECP levels (333). Normal values vary between populations, and there is conflicting information on the effect of smoking, atopy, and age (334–336). With repeated measurements, the within-subject standard deviation in log units is 0.161.

Serum ECP increases with allergen exposure or after laboratory allergen challenge and decreases after allergen avoidance and ICS therapy, although it may be less responsive than sputum eosinophil count or FeNO (337, 338). However, compared with eosinophil counts, ECP measurements in either induced sputum (339) or serum (339, 340) fail to reflect treatment-related changes in chronic asthma, suggesting that serum ECP is not a sensitive or reliable means of evaluating eosinophilic airway inflammation. Moreover, serum ECP does not appear to predict a response to corticosteroid therapy (341). Finally, a randomized trial comparing a serum ECP-based algorithm for managing asthma with a conventional algorithm found no improvement in symptom scores, in spite of increased doses of ICS (342).

<table>
<thead>
<tr>
<th>KEY POINTS AND RECOMMENDATIONS: BIOMARKERS IN CLINICAL TRIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>The role of biomarkers in asthma includes defining the phenotype at baseline, assessing underlying disease activity on treatment, and predicting the risk of future events.</td>
</tr>
<tr>
<td><strong>Induced Sputum</strong></td>
</tr>
<tr>
<td>1. Sputum induction is feasible and safe, and the techniques of sputum induction and processing have been well validated, although some technical expertise is required.</td>
</tr>
<tr>
<td>2. Assessment of eosinophilic airway inflammation using induced sputum provides additional, clinically important information about ICS responsiveness and preventable future risk of exacerbations.</td>
</tr>
<tr>
<td>3. Minimization of eosinophilic airway inflammation should be considered as an additional criterion for control of the underlying disease activity and for reduction of future risk, especially in patients with more severe asthma. Eosinophilic airway inflammation should be assessed, where possible, in clinical trials involving this population.</td>
</tr>
<tr>
<td><strong>FeNO</strong></td>
</tr>
<tr>
<td>1. FeNO measurements provide easily obtained information on underlying disease activity where it is characterized by eosinophilic airway inflammation, but the positive and negative predictive values for eosinophilia are suboptimal.</td>
</tr>
<tr>
<td>2. FeNO does not provide information about other types of airway inflammation, and this may be a problem in more severe asthma, where neutrophilic inflammation may be more important.</td>
</tr>
<tr>
<td>3. The clinical utility of FeNO-based management strategies has not been explored extensively. Currently available evidence suggests a role in identifying the phenotype in airways disease, particularly in the identification of corticosteroid responsiveness.</td>
</tr>
<tr>
<td>4. Due to logistic and cost issues, FeNO is the only biomarker likely to have a role in primary care–based asthma studies, although it is possible that with technological improvements, other techniques including sputum induction could have a role in the medium term.</td>
</tr>
</tbody>
</table>
KEY POINTS AND RECOMMENDATIONS:

CLINICAL PRACTICE

1. Where possible, biomarkers should be employed to provide information about underlying airway inflammation, a domain of the asthma “syndrome” that would not otherwise be available to the clinician.

2. Induced sputum analysis provides information on the pathology of asthma that aids decision making both as to diagnosis and treatment. The use of induced sputum to augment the clinical assessment of patients with moderate to severe asthma has been shown to be cost efficient in a specialist setting.

3. FENO measurements may be used as a surrogate marker for eosinophilic airway inflammation. They may be used to evaluate the potential for response to corticosteroid treatment.

4. Low values of FENO (< 25 ppb in adults, < 20 ppb in children) may be of particular value in aiding decisions about reducing corticosteroid dose, or alternatively for determining that ongoing airway symptoms are unlikely to be due to eosinophilic airway inflammation.

PEdiATRIC ISSUES

1. Experience with biomarkers in childhood asthma is limited, but biomarkers could prove to be useful in making an asthma diagnosis and for selecting appropriate medications based on phenotype.

2. FENO is a prototype for the application of biomarkers in children with asthma, and may be helpful in decisions on starting and stopping ICS, and perhaps monitoring medication effects.

3. Reliable measurement of FENO using the recommended single-breath online technique is limited to children 5 years and older.

4. Successful sputum induction in children is limited to those children 8 years and older, in whom success rates are around 60 to 70% in academic settings. Serial assessment of sputum may be problematic, as many children are unwilling to undergo repeat sputum inductions during follow-up visits.

RESEARCH QUESTIONS

1. More research is needed to establish whether sputum processing and analysis can be simplified to enable wider use in clinical trials.

2. More information is needed on the utility and effectiveness of sputum eosinophil-directed management in less severe asthma.

3. Information on what constitutes a clinically relevant change in sputum neutrophils is required.

4. More information is required on the utility of FENO measurement as a tool for monitoring asthma control.

5. There is a need for translational research to clarify the relationship between biomarkers and other parameters of asthma control, to establish the optimal frequency of monitoring, and to confirm the clinical and cost effectiveness of biomarker measurements in primary care and other settings.

6. More work is needed on the validation of the various measures from EBC, and to describe the relationship between these measures and other markers of asthma control. The application of non–hypothesis-driven “metabolomics” may assist in this process. Studies to address whether using EBC results in improved clinical decision-making or better asthma outcomes are required.

INDIRECT MEASURES OF ASTHMA CONTROL

Loss of asthma control potentially leads to unscheduled use of health care, loss of work and school productivity, and need for additional medication. Unscheduled use of health care may range from primary care consultations through to hospitalization or admission to an intensive therapy unit. Such episodes have significant implications for individuals and health care providers, and constitute a significant health economic burden. Measures of health care utilization provide surrogate measures for asthma control, which are particularly useful when direct clinical measures are not available, for example, at a population level. The most extreme indirect measure of poor asthma control and exacerbations is mortality, but this is more suited as an endpoint in the analysis of administrative datasets (e.g., from health maintenance organizations) than for clinical trials. Establishing the cause of death is often difficult, particularly in elderly patients.

Most existing studies have failed to define any standard methodology of reporting and there are limited data available regarding the reproducibility, responsiveness, or associations of these outcome measures.

The report and recommendations on indirect measures of asthma control are divided into four sections: (1) primary care consultations; (2) urgent health care, hospitalizations, and ER visits; (3) corticosteroid tablet usage; and (4) health economic outcomes.

Levels of Health Care

In reporting health care usage as an indirect measure of asthma control, it is important to define usage clearly, particularly for economic evaluations. This involves recognition of the fact that health care systems in different countries often define and structure primary and secondary care provision in different ways. Primary health care (i.e., initial or basic care, to which patients have direct access) may be provided by a generalist, a specialist, or a trained nurse practitioner (343). The concept of “specialists” differs between health care systems. In many European countries, for example, the specialist works in secondary care and the family doctor/general practitioner works in primary care. In other countries, specialists may work in community-based facilities but not in hospitals. In the present document, “primary care” is used to refer to clinic-based consultations to which the patient has direct access, and “secondary care” refers to visits to a hospital, ER, or equivalent facility (i.e., the classification is based on the facility at which care is provided, not the training of the health care professional).

Primary Care Consultations

Primary care consultations have been reported as an outcome measure for asthma control. There is no current standard methodology or recommendation for primary care consultations. Ways in which the variable has been reported in the past include:

- All primary care asthma-related consultations per unit time (344-347)
Some primary care consultations reflect optimal asthma management, not just for repeat prescriptions, but with evidence favoring structured proactive review at regular intervals over opportunistic or unscheduled review (354, 355). The ideal frequency of review for each patient depends on his/her disease severity and control. Hence, in publications, there is a need to distinguish between routine scheduled care consultations (or routine study visits) and unscheduled consultations, with only the latter acting as a marker of poor asthma control.

Sometimes, consultations have been subdivided according to the type of health care provider, the mode of consultation, or who initiated the consultation:

- Primary care physician face-to-face contact (345, 347, 356, 357)
- Primary care nurse (or other health care provider) face-to-face contacts (349, 358, 359)
- Telephone advice (360)
- Home visits (348)
- Doctor/practice-initiated “routine” consultations (361)
- Patient-initiated consultations (348, 349, 353, 361, 362)

Nurse-led care and telephone consultations, while common in some countries, would be unknown in other countries. Home visits because of asthma are becoming infrequent in all health care systems, and probably do not need to be reported separately.

With considerable variation in routine documentation procedures between practices, even within the same country, it is often difficult to separate planned or scheduled consultations from those that are unplanned, unscheduled, or truly emergency consultations. The definition of what constitutes an unscheduled or emergency consultation is often not specified in publications, and may range from administrative definitions (e.g., request for an appointment within 24 h) to clinical definitions (e.g., consultation judged as being needed for worsening asthma). There may be differences between the patient’s and clinician’s opinion about the reason for a consultation, especially for patients with concurrent conditions (363). It is recommended that where possible a clinical definition should be used (i.e., an unscheduled patient-initiated contact with a health care professional resulting from worsening asthma symptoms). If this definition cannot be used, the reasons should be explained, and the alternative definition justified.

Some studies have reported the number/percentage of patients requiring an unscheduled asthma appointment over a time span (e.g., 1 yr), and others, the number of consultations. It is important to distinguish between a single prolonged episode of poor control/exacerbation and multiple episodes. Where accurate information on the frequency and date of unscheduled consultations exists, this way of presenting data will give complementary information to the total/mean/median number of consultations in the studied populations, and both should be reported.

**Source of data about unscheduled consultations.** Different studies have used different ways to collect primary care consultation data; these include “unscheduled visit” trial *pro forma* completed at the time of consultation by the physician, retrospective questionnaires filled in by patients (41, 350, 364) or by physicians (361), and written or electronic clinical or administrative databases (344, 345, 365, 366). The results from different sources may not be identical, particularly as responses by patients are usually based on their recollections rather than on contemporaneous notes (363).

In order of desirability and reliability, the following methods of data collection about unscheduled care can be used for clinical trials (references are examples):

1. Standardized physician-completed data collection forms, recording symptoms, physiologic measurements, and therapy changes at the time of patient contact. This is the recommended method to provide the most robust information. If this method is not used, the reason should be explained and justified.
2. Subject-completed data collection forms, completed near to the time of contact (< 3 mo) (27)
3. Retrospective physician-completed data collection forms (361)
4. Retrospective subject-completed data collection forms. Data should be verified or corroborated from other sources if possible.
5. Inspection of routinely collected medical or administrative records (344, 345, 366, 367). This method is appropriate for pragmatic trials and observational research.

In clinical trial reports, the type, source, and definitions of visit data should be described (e.g., in an online supplement).

**Unscheduled Use of Secondary Health Care**

Good asthma control should normally be associated with no unscheduled need for secondary health care (i.e., in a hospital, ER, or equivalent facility). Because of its high cost, secondary health care has major implications for health economic outcomes. However, studies focusing on secondary health care in asthma require large sample sizes and a long duration, hence it is often not feasible to use this variable as the primary endpoint.

Reported rates of secondary health care utilization have varied, and are clearly more frequent in populations with more severe asthma. In studies of mild-moderate asthma, secondary health care visits ranged from 0.2 to 0.5/patient/year (32, 48, 246, 368–370), with higher levels (0.86/patient/year) in more severe asthma (371). It should be stated whether ER visits that result in a hospitalization are distinct from the total count of ER visits. However, the threshold for attendance in the ER or for admission to hospital will vary from country to country, reflecting global differences in the practice of medicine. Although many individual ER attendances for asthma will represent severe asthma exacerbations, as reflected by the need for systemic corticosteroids, some may represent attendances for “sick care.” There is value in recording secondary health care attendances as an overall marker of health care utilization for poor asthma control, particularly when the study focus is at a population or community level.

The recommended method of reporting results is the number of events per patient per year, expressed as the weighted mean. Data should be reported separately for ER visits, hospitalizations, and intensive care unit admissions. It is recommended that visits that occur 7 days or less from another visit should be considered to be part of the same episode. The absolute magnitude of changes in secondary health care utilization is generally small, as the events occur infrequently. Conversion of the rate to number of events per 100 patient-years may not appropriately give the perception of large changes (368).

Recording of data for unscheduled use of secondary health care is relatively easy. Since the events are infrequent, data can be collected by patient report, and/or administrative records (363). If there is a code for asthma in the first or second listed...
diagnosis, it is highly likely that the admission was at least partly, if not mostly, related to asthma.

Systemic Corticosteroid Usage

Worsening asthma control or exacerbations may require the use of additional or emergency medication. In clinical practice, this may involve several agents including bronchodilators, but the present comments are restricted to systemic corticosteroid usage (usually tablets for adults, but also including intravenous or intramuscular usage and liquid preparations in young children) as an indirect measure of asthma control.

Use of systemic corticosteroids includes both maintenance treatment in patients with severe asthma, and short courses used in the management of poor asthma control or exacerbations. Corticosteroids may be initiated either by investigators, using protocolized criteria or clinical discretion, or by patients using physician-prescribed self-management plans. There is no current standard method of reporting systemic corticosteroid usage and no documented “normal” range. In a small minority of patients with severe asthma, corticosteroid tablets may be used on a regular basis in day-to-day treatment. More commonly, corticosteroid tablets are given in short courses to treat worsening asthma. Regular corticosteroid tablet usage may therefore reflect asthma severity (a need for the regular use of corticosteroid tablets to achieve symptom control and normal lung function, i.e., “difficult to treat” asthma [6]), whereas the frequency with which courses of corticosteroids are needed is a marker of control. The Task Force Working Group on exacerbations has specified 3 days or more of systemic corticosteroid use as a mandatory criterion for severe exacerbations, so data for systemic corticosteroid use will closely parallel data for severe exacerbations, but will also include early cessation by patients of a longer prescribed course (372) or patient self-administration for shorter periods of time.

The way in which short-term use of systemic corticosteroids has been recorded in clinical trial reports is far from uniform. Some reports have recorded the number of patients needing oral or systemic corticosteroids as a marker of loss of control (373) and others have used a need for systemic corticosteroids to define a serious asthma-related attack without quantifying severity in any other way (41, 364, 375). Of the studies which were reviewed, many were too short to report systemic corticosteroid use as an outcome (365, 366). When reported, systemic corticosteroid use was most commonly expressed as mean number of courses per unit time, usually per year (27), with some studies also reporting the percentage of patients needing a course of corticosteroids (367), or days with corticosteroid tablets due to exacerbations (32).

One potential problem with reporting systemic corticosteroids relates to the lack of quantification of corticosteroid dosage. The most accurate measurement of corticosteroid use would be to report the milligrams of prednisolone taken per patient per unit time. However, trial reports do not always make it clear whether the corticosteroid dose was standardized in the protocol or left to physician discretion, and the duration of a course is rarely mentioned, nor is the handling of closely consecutive courses described. Daily corticosteroid dose would also be difficult to calculate if some patients were on regular maintenance therapy, where it would need to be recorded as extra milligrams of prednisolone per patient per unit time.

Accordingly, in the absence of definite evidence that the quantity of corticosteroid taken is a better marker of control than the need for any corticosteroid, it is recommended that the total number of courses of corticosteroids per patient per year should be recorded, but that all studies should record whether courses were standardized or left to individual physician’s discretion. In those taking maintenance oral corticosteroids, a “course” is defined as a short-term increase in dose with a subsequent reduction to the baseline level. Further study of the relationship between corticosteroid bursts and hospitalization is needed.

Health Economic Data

In general, loss of asthma control and exacerbations lead to higher medical and nonmedical costs and lower quality of life. A full and detailed accounting of quantifiable costs is required to reflect the economic impact of poor control and of therapeutic interventions. Health economic data are frequently collected to support local or national coverage and reimbursement of interventions. They comprise data about the direct cost of medical resource utilization (e.g., health care contacts, hospital use, and medications) as well as data about indirect costs, established from patient-reported data concerning loss of work/school time and quality of life. Data regarding school or work absence has been used as an endpoint in some studies (usually of an educational rather than a pharmacologic intervention), but it is likely to be a poor reflection of control. The Work Productivity Assessment Instrument (WPAI) has been developed and validated as a tool for characterizing the degree of absence and productivity impairment in asthma for health economics and burden estimation, but the WPAI is only useful for characterization, not for outcome evaluation. Reilly and colleagues (375) were the first to report validity and reproducibility of the WPAI; more recently, the WPAI has been used as a measure of disease burden in asthma (376).

Resource utilization data from clinical and observational studies are used to create aggregated patient-level profiles of health care utilization. It is important to identify the dates of service so that unique episodes of care (which may include several health care visits) can be constructed. Utilization data are then combined with local and relevant health care unit price data to give an estimate of health care costs. Unit price data should not be collected as part of a clinical trial protocol. Rather, a separate and parallel protocol should be established for the purpose of ascertaining nationally representative (or jurisdiction-specific) price information.

Typically, health care cost data are reported in terms of weighted mean or median costs per patient per year. The main findings from economic evaluations should be reported in local currency, as this is more relevant to local decision makers. In addition, a more global currency (USD, EURO) can be provided, particularly for multinational studies or for publications in international journals.

### Key Points and Recommendations: Clinical Trials

1. Primary care consultations should be expressed as the weighted mean or median rate per patient per year, divided into scheduled and unscheduled consultations. Only unscheduled consultations serve as a marker of poor asthma control; these should be defined as those initiated by the patient because of worsening asthma. Data should be collected, where possible, from standardized physician-completed forms at the time of patient contact. Telephone consultations and nurse consultations should be reported separately.

2. Unscheduled use of secondary health care is most useful as an outcome measure in populations in which such
The above statements are also relevant to clinical trials conducted in primary care.

### SUMMARY FOR CLINICAL PRACTICE

1. Primary care consultations should be recorded during the visit as scheduled or unscheduled, based on a clinical assessment, and further classified by mode (office visits, home visits, telephone consultations etc.). The record should state who saw the patient (doctor, nurse, educator, etc.)

2. All hospital referrals of patients with asthma, and the outcome, should be recorded. Follow-up arrangements should be documented so that the patient can be recalled to see the primary care doctor if necessary.

3. All prescriptions for systemic corticosteroids for asthma exacerbations should be recorded. Patients prescribed these medications in reserve (for later use with self-management plans) should be provided with means to record when and why they have used the medications.

4. Copies of medical certificates issued for absence from work due to asthma should be recorded in the notes.

### PEDIATRIC ISSUES

1. Estimating unscheduled health care consultations for asthma in children is complicated by consultations for infection-related nonspecific airway symptoms.

2. Antibiotics and short courses of high-dose ICS are frequently, though often inappropriately, prescribed for asthma exacerbations in children. They should be considered when monitoring respiratory events possibly related to asthma, especially in young children.

3. Assessment of nonmedical costs of poor asthma control should consider not only the child’s absence from school, but also loss of parental work-related productivity.

4. Determination of health care utilization and nonmedical costs in pediatric and adolescent clinical trials is best based on reports by parents or caregivers rather than patients.

### RESEARCH QUESTIONS

1. How and why patients and physicians decide to use/increase systemic corticosteroids is unclear, and how such use is recorded varies. Further work is needed to clarify whether any objective measures predict need/increase of corticosteroids and whether a more detailed recording (for example mg prednisolone/patient/unit time) enhances understanding or quantification of control.

2. As regards time off school, and consequent parental costs, more work is needed to clarify what influences need for school absence, which may not always reflect loss of asthma control alone (i.e., it may also be influenced by other socioeconomic factors).

3. More work is needed to understand how health economic evaluations can take into account the existing variation in delivery between different countries or regions.

### HEALTH-RELATED QUALITY OF LIFE (HRQOL)

For over a decade, generic or asthma-specific quality-of-life questionnaires have been included in clinical trials to assess treatment benefits as perceived by the patient. They have been used because (1) some treatment effects can be only identified by the patient, (2) patients provide a unique global perspective on treatment effectiveness, and (3) standardized assessment may be more reliable than informal interview.

Quality of life has been defined as “the functional effects of an illness and its consequent therapy upon a patient, as perceived by the patient” (377). It includes somatic sensation (the problems associated with symptoms), physical and occupational function, emotional and psychological impact, and social interaction. Factors contributing to a sense of well-being include good health, a secure social and occupational environment, financial security, spirituality, self confidence, and strong, supportive family relationships. Health-related quality of life can be considered as the component related to the overall burden of a chronic disease with respect to these domains. It is important to use specifically designed tools to assess health-related quality of life (HRQOL), and to differentiate the information that they provide about treatment response from other subjective beliefs and clinical outcomes.

**Why Measure HRQOL in Asthma?**

Most studies using generic or disease-specific instruments have reported that asthma affects HRQOL (378, 379), with lower quality-of-life scores being found in patients with more severe asthma (380). However, clinical trials in asthma have often focused on outcomes that are primarily of importance to the clinician, such as symptom scores or lung function. These do not necessarily reflect all the characteristics of the disease. For example, there are no data reporting the relationship between inflammatory markers and HRQOL. The patient’s perception of the burden of disease may be completely different from the clinician’s, and may vary according to the patient’s circumstances and life expectations. Measuring HRQOL can add...
valuable information to better assess the impact of poor asthma control and/or its severity (i.e., difficulty to treat [6]). Initially, HRQOL questionnaires were not intended to be used as endpoints in clinical trials, but many studies now include an assessment of HRQOL in the assessment of effects of therapy, both pharmacologic (52, 381, 382) or nonpharmacologic (383), to assess the global benefit to patients.

**Interpretation of HRQOL Results**

The application of HRQOL questionnaires is far from established in clinical practice, and there are no clear guidelines for their use in the assessment of asthma control. HRQOL data should be interpreted with caution, as the performance characteristics of an instrument (score, validity, reproducibility, and responsiveness) may vary from the original description and purpose according to the context of the study. For most of the asthma-specific quality of life questionnaires, an estimation has been made of the MID, the smallest difference in score that patients perceive as beneficial and that, in the absence of troublesome side effects and excessive cost, would mandate a change in a patient’s management (384–387). Importantly, MID for quality of life has often been assessed from the investigator’s perspective, but is now primarily anchored to the patient’s perspective, on the basis of the idea that that only the patient is in a position to judge whether a difference is important.

**Choice of HRQOL Questionnaires**

It is a difficult task to select the best HRQOL tool to be used in a clinical routine setting or in a clinical trial. One should select the most adequate, easy to use, best validated, and responsive questionnaires. Use of some questionnaires is subject to copyright restrictions and charges. The ATS website provides a comprehensive list of English-language generic and asthma-specific HRQOL tools for adults and children (388). Several well-validated HRQOL tools are described below and in Table E4. All were developed using broadly similar methodologies and have a similar general structure and content. There are two types of quality of life questionnaires: generic and specific.

**Generic HRQOL Questionnaires**

Generic HRQOL questionnaires such as the SF-36 (389, 390) were designed for use by patients with any chronic illness, and they include items such as tiredness, headache, or gastric problems. Their utility is questionable in the context of asthma. The real advantage of these instruments is to compare the burden of different chronic conditions (e.g., asthma and arthritis). By covering a wide spectrum of symptoms and activities, they are too superficial to reflect the reality of a patient's life with asthma, and they should be complemented by a more specific tool. Generic instruments are limited by their inability to identify specific problems in individual patients and their lack of responsiveness to small but potentially important changes in quality of life. These views have been challenged and at least one study showed that the SF-36 instrument was able to detect changes in asthma “severity” as well as treatment effects (391, 392). Generic health-related QOL instruments have also been developed and validated for children of various ages (393, 394).

**Measurement methods and content.** In adults, the most commonly used and the best-validated generic HRQOL tool is the SF-36 (389, 390). It includes 36 items measuring three major health attributes and nine health concepts. It is self-administered and can be completed in 10 minutes. This tool has been validated and found to be reliable, with translations available in many languages.

Reference values, reproducibility, and responsiveness. Scoring of the SF-36 is complex, and yields eight scale scores ranging between 0 and 100. A high score is consistent with a positive health status. The SF-36 has good internal consistency and cross-sectional validity in patients with asthma (391).

**Comparison with other disease-related outcomes.** Many studies have reported low to moderate relationships between airflow limitation (measured by FEV1), respiratory symptoms, and HRQOL. In one cross-sectional study of patients with a broad range of asthma severity, rho for correlations between FEV1 and SF-36 scales ranged from 0.09 for mental health to 0.40 for physical functioning (395).

**Specific Asthma-related QOL questionnaires**

Currently, there are several validated asthma-related quality of life questionnaires. These instruments include the functional impairments (physical, emotional, social, and occupational) that are most important to patients with asthma. They have been reported to be much more sensitive to change in patients’ quality of life than generic health profiles. Questionnaires with substantial validation data are summarized below, and more detail is provided in Table E4.

**AQLQ**

The standardized version of the AQLQ is a 32-item, disease-specific questionnaire that is reported as having strong measurement properties and validity for measurements of functional impairment in adults with asthma (396, 397). Patients score their experiences during the last 2 weeks on a 7-point scale (1 = severe impairment to 7 = no impairment). The overall AQLQ score and mean responses for different domains (symptoms, activities, emotions, and environment) are calculated. The MID is reported to be 0.5 points (383). The AQLQ has proved responsive in before-after studies and in clinical trials (382, 398, 399). Construct validity has been assessed using both conventional measures of asthma severity and generic quality of life instruments. The AQLQ has been shown to correlate with asthma control questionnaires such as ACQ (158) and ATAQ (226). There is a poor correlation between change in AQLQ and change in FEV1 (399).

For pediatric studies, Juniper and colleagues have also validated the Pediatric Asthma QOL Questionnaire (PAQLQ) (400) and the Pediatric Asthma Caregivers QOL Questionnaire (PACQLQ) (401).

**Mini-AQLQ**

This questionnaire is a short version of the AQLQ, and includes five items on symptoms, four items on activity limitations, three on emotional function, and three concerning environmental stimuli, scored with the same 7-point scale as the AQLQ (402). Measurement properties of the Mini-AQLQ are good, but not as strong as the original AQLQ.

**AQOL**

This is a 20-item self-administered questionnaire (403, 404) that provides a total score together with subscale scores for breathlessness, mood disturbance, social disruption, and concerns for health, calculated by averaging of item scores. Lower scores indicate better quality of life. Validation studies showed good short-term test-retest reproducibility. The questionnaire was internally consistent in a sample of outpatients and in a community sample with asthma. Weak correlations in the expected direction were seen with three markers of asthma severity (spirometry, AHR, and number of asthma medications) (403).

The validity and responsiveness of AQOL were assessed in 44 adults with asthma (404). Change in AQOL score was significantly correlated with change in symptom score and change in AHR, with only a trend for change in peak flow variability and in Sickness Impact Profile score. The AQOL was capable of detecting differences between improved and stable subjects ($P = 0.007$).
Adams and colleagues modified the AQLQ Marks to a 22-item questionnaire using a 7-point Likert response scale (MAQLQ-M) (406), inverted so that higher values represented better quality of life. Adult subjects with moderate-severe asthma were evaluated at baseline and 3-month follow-up. In cross-sectional and longitudinal analysis, stronger associations were seen with symptom and self-rating scales than with lung function, medication usage, or health service utilization measures. Higher baseline scores were associated with lower risks over 12 months for hospital admissions and repeated ER visits.

The Living with Asthma Questionnaire (LWAQ). The LWAQ (407) has 68 items covering 11 domains of asthma experience, which were derived from focus group discussions. The scale compensates for acquisition bias as well as allowing a “not applicable” response category. Validity of the scale was demonstrated by confirmation of expected group differences, and the retest reproducibility was 0.948. A shorter version of the LWAQ has been published (389).

The St. George’s Respiratory Questionnaire. The SGRQ (408) is a 76-item self-completed questionnaire that was developed for assessment of quality of life in COPD. It is widely used in COPD clinical trials (409), but has also been validated and used in asthma (408, 410, 411).

The Asthma Questionnaire-20 (AQ20). The AQ20 (412) was developed as a short and simple measure of health status in asthma, using 20 dichotomous responses (yes/no) relating to the effects of asthma on the patient’s life, emotions, and activities. The reproducibility of the AQ20 is high (413). The AQ20 score correlates with AQLQ Juniper \( r = 0.40, P < 0.001 \) and with SGRQ \( r = 0.46, P < 0.0001 \) (414).

Role of HRQOL Assessment in Drug Evaluation Process

At present, the role of HRQOL assessment in clinical trials is linked to the willingness of the study sponsor to incorporate such measures into the process of drug development. Regulatory agencies such as EMEA and FDA have shown interest in patient-reported outcomes and specifically in HRQOL, and the FDA has provided draft guidelines to facilitate the development of robust questionnaires and to better understand the information obtained from HRQOL.

Several large-scale clinical asthma trials have incorporated the AQLQ Juniper as a measure of the global impact of asthma control (13, 415, 416). In the Formoterol and Corticosteroids Establishing Therapy (FACET) study, the correlation in individual patients between changes in clinical indices and changes in AQLQ score during the 12-month randomized period were weak to moderate (maximum \( r = 0.51 \)) (13).

In conclusion, HRQOL questionnaires measure the impact of asthma on the individual, and provide complementary rather than direct information about asthma control or severity. Several HRQOL instruments have been developed and validated, with an estimate of an MID in HRQOL score provided for some of these. Further work is needed to determine the specific role of HRQOL measures in clinical trials and drug development.

KEY POINTS AND RECOMMENDATIONS: CLINICAL TRIALS

1. HRQOL is a patient-reported outcome that represents the overall impact of the level of asthma control and exacerbations on quality of life. It should be used as a specific assessment tool in asthma clinical trials.

2. To correctly interpret changes in HRQOL, the MID should be defined for each validated HRQOL measure.

3. Potential gains in HRQOL resulting from treatment may be offset by the impact of drug-related side-effects or co-morbidities.

4. The effect of cultural and educational differences on HRQOL assessment should be considered in the development and use of questionnaires.

5. Generic health-related QOL instruments have been validated and may be used to compare the impact of asthma with that of other chronic illnesses.

6. Copyright and conditions of use should be checked before quality of life questionnaires are used in clinical trials.

Each of these points is also applicable for clinical trials in primary care.

KEY POINTS AND RECOMMENDATIONS: PEDIATRIC ISSUES

1. All disease-specific quality-of-life instruments used for pediatric studies should be validated for relevant age groups.

2. The Pediatric Asthma Quality of Life Questionnaire (PAQLQ) is a pediatric disease-specific QOL instrument designed and validated in several languages for children aged 6 to 18 years. An asthma-related QOL questionnaire for caregivers has been developed by the same authors.

3. For any pediatric QOL instrument, the study protocol should specify whether the carer or the child should answer the QOL measure.

4. Child-completed QOL questionnaires must take into consideration the child’s reading level. Children under 12 years of age may have difficulty reading or understanding a questionnaire without assistance.

5. When children are assisted by their parent in completing a questionnaire, their responses change. Therefore, child-completed questionnaires should either be completed by the child alone or with the assistance of professional staff—as specified in the protocol for that test.

RESEARCH QUESTIONS

1. The comparative usefulness of validated disease-specific HRQOL questionnaires and composite asthma control scores or other asthma-related outcomes should be evaluated in clinical trials and clinical practice.
2. The time-course of change in HRQOL, and the numbers of required assessments, should be better defined. Long-term trials are recommended due to the potential of exacerbations to affect HRQOL in asthma.

3. The relationship between HRQOL and future risk of adverse events should be examined, as HRQOL may be influenced by factors that drive health-related behavior such as medication adherence.

4. The positioning of HRQOL questionnaires within the approval process for new therapeutic interventions by regulatory bodies should be further evaluated.

5. There is a need for clinical trials to assess whether the use of formal HRQOL tools in routine clinical care leads to better outcomes.

### SUMMARY AND OVERALL RECOMMENDATIONS

The recommendations of the Task Force are based on the development of a model, published elsewhere (6), that links the concepts of asthma phenotypes, underlying disease activity, asthma severity, and asthma control. “Asthma control” is defined as the extent to which the manifestations of asthma have been reduced or removed by treatment. It should be assessed not only by current clinical features such as symptoms, reliever use, and lung function, but also by evaluation of the patient's risk of adverse outcomes (e.g., exacerbations or medication side-effects) in the future. “Asthma severity” is now defined as the intensity of treatment required to achieve good asthma control; for severe asthma, there is a requirement for (not necessarily just prescription or use of) high-intensity treatment, and mild asthma can be well controlled with low-intensity treatment. Asthma severity may be influenced by the underlying disease activity, and by the patient’s phenotype, both of which may be further described using pathologic and physiologic markers. These markers may also act as surrogate measures for future risk. Biomarkers provide a link between phenotype, severity, and control (6).

### Table 1. Recommended Outcome Measures Relating to Asthma Control for Clinical Trials: Rationale

<table>
<thead>
<tr>
<th>Definitions</th>
<th>Method of Observation</th>
<th>Purpose of Recording the Outcome Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baseline characteristics</td>
<td>• Direct measurement at baseline</td>
<td>• To describe the patients’ pretreatment level of asthma control</td>
</tr>
<tr>
<td>• Characteristics of the study population before randomization</td>
<td>• Baseline – randomization visit or (for diary measures) the final 2 wk of run-in</td>
<td>• To characterize the study population in terms of asthma phenotype(s) and underlying disease activity</td>
</tr>
<tr>
<td>• To record baseline levels of predictors* of future risk</td>
<td>• To record baseline levels of predictors* of future risk</td>
<td>• To identify discrepancies between current clinical control and markers of underlying disease activity, e.g., with masking by LABA monotherapy or in poor perceivers</td>
</tr>
<tr>
<td>2. Outcome measures for the assessment of treatment effect . . .</td>
<td>• Direct measurement of level of current clinical asthma control throughout the study⁴</td>
<td>• To assess the effect of the treatment on level of clinical control</td>
</tr>
<tr>
<td>... on current clinical control</td>
<td>• On each occasion, level of control should be assessed over the previous 1-4 wk</td>
<td>• To describe the range of responses to study treatment</td>
</tr>
<tr>
<td>• To characterize the study population</td>
<td>• To identify discrepancies between current clinical control and markers of underlying disease activity, e.g., with masking by LABA monotherapy or in poor perceivers</td>
<td></td>
</tr>
<tr>
<td>... on future risk</td>
<td>• Future risk refers to the risk of adverse outcomes such as exacerbations, poor asthma control, or accelerated decline in lung function, or side-effects of treatment, in the near or distant future</td>
<td>• To assess the effect of the treatment in reducing the occurrence of those adverse outcomes which can be directly measured in the study, e.g., in a long-term clinical trial, the number of exacerbations in the treatment and control groups can be directly recorded</td>
</tr>
<tr>
<td>• Future risk refers to the risk of adverse outcomes such as exacerbations, poor asthma control, or accelerated decline in lung function, or side-effects of treatment, in the near or distant future</td>
<td>• Direct measurement, e.g., number of exacerbations, or decline in post-BD FEV₁, during a long-term study⁵</td>
<td>• To predict the effect of the treatment in reducing those adverse outcomes which are not able to be quantified in this particular study (e.g., because the study is too short or it is not powered for major adverse outcomes), by recording the effect of the treatment on predictors of future risk.*</td>
</tr>
<tr>
<td>• Indirect assessment based on probability—the extent to which the treatment leads to improvement from baseline in predictors* of future risk (surrogate measures).</td>
<td>• To assess the effect of the treatment in reducing the occurrence of those adverse outcomes which can be directly measured in the study, e.g., in a long-term clinical trial, the number of exacerbations in the treatment and control groups can be directly recorded</td>
<td></td>
</tr>
</tbody>
</table>

*Definition of abbreviations: BD = bronchodilator; LABA = long-acting β₂-agonist.

*“Predictors” are modifiable factors that have been observed to be associated with increased risk of adverse asthma outcomes in the future (such as exacerbations, future poor asthma control, and accelerated decline in lung function, or side effects of treatment), and that can be used as surrogate measures in studies in which the adverse outcomes cannot be directly measured. An example of a predictor is F₅₀, which has been observed to be associated with an increased risk of exacerbations.

¹ In general, it is preferable to record the level of asthma control throughout the study rather than just at the end of the treatment period. Use of data from the whole treatment period will reflect the magnitude and rate of treatment response, the extent of variation in level of control, and the occurrence of exacerbations, all of which are relevant to the overall impact of treatment on the patient. This will typically include assessment at each study visit, or, for diary measures, over multiple periods each of 1–4 wk. The resulting multiple data points for each patient can be analyzed by mixed model (or equivalent) analyses, which are to be preferred over merely averaging the data over the whole treatment period.
### TABLE 2. RECOMMENDED OUTCOME MEASURES RELATING TO ASTHMA CONTROL FOR CLINICAL TRIALS: LIST OF MEASURES

<table>
<thead>
<tr>
<th>Minimum Set of Measures (Essential)</th>
<th>Desirable</th>
<th>Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baseline characteristics*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Symptom-free days†</td>
<td>- Airway hyperresponsiveness‡</td>
<td></td>
</tr>
<tr>
<td>- Reliever use‡</td>
<td>- Biomarkers**</td>
<td></td>
</tr>
<tr>
<td>- Pre-BD FEV₁§</td>
<td>- Treatment side-effects§</td>
<td></td>
</tr>
<tr>
<td>- Post-BD FEV₁†</td>
<td>- History of exacerbations</td>
<td></td>
</tr>
<tr>
<td>- Composite scores§</td>
<td>- (OCS, ER visits, hospitalizations)**</td>
<td></td>
</tr>
<tr>
<td>- Quality of life†</td>
<td>- On-treatment FEVᵢ,ili</td>
<td></td>
</tr>
<tr>
<td>2. Outcome measures for the assessment of treatment effect . . .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>. . . on current clinical control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Symptom-free days†</td>
<td>- On-treatment FEVᵢ,ili</td>
<td></td>
</tr>
<tr>
<td>- Reliever use‡</td>
<td>- Symptom/reliever/lung function</td>
<td></td>
</tr>
<tr>
<td>- Composite scores§</td>
<td>- diary**</td>
<td></td>
</tr>
<tr>
<td>- Exacerbation (within last 1-4 wk)**</td>
<td>- Indirect measures, e.g., corticosteroid use, health care utilization</td>
<td></td>
</tr>
<tr>
<td>- Quality of life†</td>
<td>- Airway hyperresponsiveness‰</td>
<td></td>
</tr>
<tr>
<td>. . . on future risk</td>
<td>- Biomarkers‰</td>
<td></td>
</tr>
<tr>
<td>For direct measurement of adverse outcomes:</td>
<td>- Airway hyperresponsiveness‰ (as predictor of future risk)</td>
<td></td>
</tr>
<tr>
<td>- Exacerbations***</td>
<td>- Biomarkers‰</td>
<td></td>
</tr>
<tr>
<td>- Post-BD FEV₁† (for assessment of lung function decline)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Composite scores§</td>
<td>- For indirect assessment of risk of adverse outcomes</td>
<td></td>
</tr>
<tr>
<td>- Treatment side-effects§</td>
<td>- Airway hyperresponsiveness‰ (as predictor of future risk)</td>
<td></td>
</tr>
<tr>
<td>For indirect assessment of risk of adverse outcomes</td>
<td>- Biomarkers‰</td>
<td></td>
</tr>
<tr>
<td>- Pre-BD FEV₁ (as predictor for exacerbations)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Definition of abbreviations:** BD – bronchodilator; ER – emergency room; OCS – oral corticosteroids.

* Baseline characteristics are assessed at the randomization visit, or, for diary measures, during the final 2 wk of run-in.
† Symptom-free days and β₂-agonist use (reliever-free days and occasions/day) may be ascertained from a diary or from a visit-based questionnaire (or from suitably-worded components of a composite score). If a visit-based questionnaire or composite score is used, the period of assessment for reliever use and symptom-free days should be no more than 4 wk. Symptom-free days are not suitable as an outcome measure for study populations with very frequent or very infrequent symptoms.
§ Post-BD FEV₁ is defined as FEV₁ recorded after appropriate withholding of short-acting and long-acting bronchodilator, if used.
†† Airway hyperresponsiveness is a marker of underlying disease activity, and the extent to which this has been modified by treatment. It allows assessment of discrepancies with the observed level of clinical control (e.g., with masking by LABA monotherapy). In the assessment of treatment effect, airway hyperresponsiveness also serves as a predictor of future risk (see footnote to previous table).
†‡ (as predictor for health care utilization, ER visits, hospitalizations); mortality due to asthma
** Diary measures should be obtained from validated diary questions and, where possible, using electronic data collection to improve data quality and avoid data fabrication. Morning PEF is the most consistently reported lung function variable from diaries.
§§ Airway hyperresponsiveness is a marker of underlying disease activity, and the extent to which this has been modified by treatment. It allows assessment of discrepancies with the observed level of clinical control (e.g., with masking by LABA monotherapy). In the assessment of treatment effect, airway hyperresponsiveness also serves as a predictor of future risk (see footnote to previous table).
** Airway hyperresponsiveness is a marker of underlying disease activity, and the extent to which this has been modified by treatment. It allows assessment of discrepancies with the observed level of clinical control (e.g., with masking by LABA monotherapy). In the assessment of treatment effect, some biomarkers also serve as predictors or surrogate measures of future risk. The storing of DNA may offer the opportunity to study gene–environment interactions that affect future risk. The approach should comply with local ethical guidelines.
†† Treatment side effects: record side effects relevant to study medication(s), as-needed medications, or exacerbation medications, and any withdrawals due to adverse events. Note that some side effects related to asthma medications (e.g., dysphonia [ICS] or mood changes [OCS]) may not be perceived by patients as “health problems” and therefore may be underestimated by routine Adverse Event questioning.
††† In most cases, the history of previous exacerbations cannot be directly compared with prospectively recorded exacerbations because of recall errors, but is important for characterizing patients at entry.
‡‡‡ On-treatment FEV₁ is defined as FEV₁ recorded without withholding of study medication. To standardize the measurement, it should be performed 6 hours after SABA where possible. “On-treatment FEV₁” is only substantially different from “Pre-BD FEV₁” for studies in which subjects are taking LABA. In such studies, preference should be given to recording “Pre-BD FEV₁,” (where LABA is withheld) because of the additional information that this measure provides about future risk.
‡‡‡‡ Severe exacerbations have been defined by the Task Force as events requiring systemic corticosteroids for > 3 d and/or a hospitalization/emergency room visit for asthma requiring systemic corticosteroids.

(e.g., ambulatory lung function monitoring is not generally appropriate in this age group). Details are given in the pediatric “boxes” within this document.

The principles described above also apply for the assessment of asthma control in clinical practice. Specific recommendations for clinical practice are found at the end of each of the sections of this document.

### Choice of Endpoints

Some therapeutic interventions principally affect clinical control, others target future risks, and others (e.g., disease modifiers) may impact both. Likewise, some measures provide information about current clinical control, some about future risk, and some about both. Hence, it is not appropriate to recommend a single endpoint for the assessment of asthma control. This is reflected in Tables 1 and 2.

In long-term therapeutic studies, the effect of the intervention on some future risks (e.g., exacerbations), may be measured directly. However, it is impractical for all studies to be of sufficient duration to directly measure these events, so, in shorter studies, inferences about some future risks may be drawn from measurement of biomarkers and physiologic measures.

From Table 2, it is recommended that all clinical trials that aim to study the effect of an intervention on asthma control should include and report the “minimum” or “essential” out-
come measures for both current clinical control and future risk. Each study report should refer to the extent to which both of these goals of treatment have been met. If it is not possible in a particular clinical trial to include any measures relevant to “future risk,” this limitation of the study design, and the implications for the conclusions from the study, should be acknowledged.

For some trials, the nature of the study hypothesis may mandate additional or alternative outcome measures. These include trials that are primarily intended to investigate the effect of treatment on the underlying disease process or those intended to evaluate specific pharmacologic effects. The choice of the primary endpoint in any individual study may vary according to the specific intervention.

Care should be taken in the interpretation of existing studies or guidelines that refer to asthma control, severity, and exacerbations, as the concepts underpinning the analysis or recommendations may be different from the new definitions established by the Task Force (6).

Future Directions

The present Task Force has provided a redefinition of the concepts of asthma control and severity (6), so that the rationale behind the assessment of asthma may be improved. The new definitions provide a framework for recommendations regarding the optimum assessment of asthma control in clinical trials. However, these are not the final answers, and the overall landscape may yet change. Further progress will result from integrated studies that variously combine the clinical, physiologic, and/or pathologic measures which, when used to guide treatment, are found to lead to the best outcomes for patients. Single outcome measures may not suffice for assessment of treatment response in the context of the current multi-component definition of asthma. The existing composite asthma control scores are a first step toward developing measurement tools that will integrate information about current clinical control and future risk, to reflect both the benefits and potential harm of treatment. Such tools need to be simple enough for use in clinical practice as well as in clinical trials. Extending this approach to include bioinformatics-style techniques in large datasets will identify the best combination and weighting of factors, recorded at baseline or during the early course of treatment, that predict a patient’s subsequent course. For example, in a recent study by Osborne and coworkers (66), the final models predicting need for acute care were very different from current composite clinical scores.

Emerging work on biomarkers suggests that several measures of airway inflammation should be evaluated in a similar way for inclusion in composite control scores. Cluster analysis has identified clinical phenotypes that are characterized by discordance between symptoms and airway inflammation, and that predict response to sputum-guided treatment (288). In relatively asymptomatic chronic diseases such as hypertension and diabetes, treatment decisions are based on biomarkers rather than on symptoms (417), but their use is not without criticism. However, biomarkers may ultimately prove to be more appropriate in asthma, particularly for some phenotypes. This approach is already accepted in asthma for patients with poor perception of airway obstruction (73). The possibility needs to be envisaged that, in the future, a well-validated biomarker might override symptoms as the basis for treatment decisions: while not ignoring the burden of symptoms to patients, we ought not to be locked in to the definitions of asthma control provided in this document, or the current stepwise approach to treatment.

This Statement was prepared by a Task Force of the Assembly on Allergy, Immunology, and Inflammation of the American Thoracic Society and the Inflammatory Airway Diseases and Clinical Allergy Assembly of the European Respiratory Society.

The members of the Task Force were:

Eric D. Bateman, M.B., Ch.B., M.D., FRCP, DCH, University of Cape Town Lung Institute, Cape Town, South Africa
Louise-Philippe Boulet, M.D., FRCP, FCCP, Institut de Cardiologie et de Pneumologie de l’Hôpital Laval, Québec, PQ, Canada
Homer A. Boushey, M.D., University of California, San Francisco, CA
William W. Busse, M.D., University of Wisconsin, Madison, WI
Thomas B. Casale, M.D., Creighton University, Omaha, NE
Pascal Chanez, M.D., Ph.D., Université de la Méditerranée, Marseille, France
Paul L. Enright, M.D., University of Arizona, Tucson, AZ
Peter G. Gibson, M.B., B.S., Centre for Asthma & Respiratory Disease and Hunter Medical Research Institute, University of Newcastle, Australia
Johan C. de Jongste, M.D., Ph.D., Sophia Children’s Hospital, Rotterdam, The Netherlands
Hui A. M. Kerstjens, M.D., Ph.D., University Medical Center, Groningen, The Netherlands
Stephen C. Lazarus, M.D., University of California, San Francisco, CA
Mark L. Levy, M.B., Ch.B., FRCP, University of Edinburgh, Edinburgh, UK
Paul M. O’Byrne, M.B., FRCP(C), Firestone Institute for Respiratory Health, McMaster University, Hamilton, ON, Canada
Martyn R. Partridge, M.D., FRCP, Imperial College, London, UK
Ian D. Pavord, D.M., FRCP, University of Leicester, Leicester, UK
Heleen K. Reddel, M.B., B.S., Ph.D., FRACP (co-chair), Woolcock Institute of Medical Research, Camperdown, Australia
Malcolm R. Sears, M.B., Ch.B., FRACP, FRCP, Firestone Institute for Respiratory Health, McMaster University, Hamilton, ON, Canada
Peter J. Sterk, M.D., Ph.D., University of Amsterdam, Amsterdam, The Netherlands
Stuart W. Stolofov, M.D., University of Nevada School of Medicine, Reno, NV
Stanley J. Szefler, M.D., National Jewish Medical and Research Center and University of Colorado Health Sciences Center, Denver, CO
Sean D. Sullivan, B.S.Pharm., M.Sc., Ph.D., University of Washington, Seattle, WA
D. Robin Taylor, M.D., MRCP (UK), FRCP (co-chair), Dunedin School of Medicine, University of Otago, Dunedin, New Zealand
Mike D. Thomas, M.B., B.S., FRCP (Lond), University of Aberdeen, Aberdeen, UK
Sally E. Wenzel, M.D., University of Pittsburgh, Pittsburgh, PA

Observers:

Eugene J. Sullivan, M.D., FCCP (FDA)
Badruel A. Chowdhury, M.D., Ph.D. (FDA)
David J. Lyons, M.D., M.Sc., FRCP, MRCP (EMEA)

The members of the Task Force Working Groups were:

Exacerbations: Paul M. O’Byrne, Thomas B. Casale, Mark L. Levy, Helen K. Reddel, Stanley J. Szefler
Diary Data: Johan C. de Jongste, Hui A. M. Kerstjens, Stephen C. Lazarus, Helen K. Reddel
Lung Function and Airway Hyperresponsiveness: Paul L. Enright, Homer A. Boushey, Peter J. Sterk
Composite Scores: Eric D. Bateman, Malcolm R. Sears, Helen K. Reddel, D. Robin Taylor
Biomarkers: Ian D. Pavord, Peter G. Gibson, D. Robin Taylor
Indirect Measures: Martyn R. Partridge, Sally E. Wenzel, Sean D. Sullivan, Mike D. Thomas


82. Reddel HK, Toelle BG, Marks GB, Jenkins CR, Woolcock AJ. Analysis of adherence to peak flow monitoring when recording of data is electronic. BMJ 2002;324:146–147.


100. Tasche MJ, van der Woonden JC, Uijen JH, Ponsioen BP, Bermsen RM, van Suijlekom-Smit LW, de Jongste JC. Randomised placebo-controlled trial of inhaled sodium cromoglycate in 1–4-year-old


186. Cockcroft DW, Davis BE. The bronchoconstrictive effect of inhaling methacholine by using total lung capacity inspirations has a marked influence on the interpretation of the test result. *J Allergy Clin Immunol* 2006;117:1244–1248.


Juniper EF, Price DR, Stampone PA, Creemers JPHM, Mol SJM, Fireman P. Clinically important improvements in asthma-specific quality of life, but no difference in conventional clinical indexes in patients changed from conventional beclomethasone dipropionate to approximately half the dose of extrafine beclomethasone dipropionate. Chest 2002;121:1824–1832.


