Clinical Outcome Measures of Specific Immunotherapy

Oliver Pfaar; Clemens Anders; Ludger Klimek
Published: 07/20/2009

Abstract and Introduction

Abstract

Purpose of Review: To provide an overview of clinical parameters generally used for monitoring the clinical efficacy of specific immunotherapy (SIT) in clinical trials. In particular, it focuses on primary and secondary outcome measurements and reviews the advantages and disadvantages of each method.

Recent Findings: In 2007, the World Allergy Organization defined the severity of symptoms and the need for concomitant medication as primary endpoint parameters in clinical outcome measures of SIT. Furthermore, it was stated that the symptom score should always be combined with the rescue medication score. The ‘quality of life’ is usually used as a secondary outcome measure in clinical trials on SIT.

Summary: In clinical trials on SIT, several clinical parameters are commonly used to provide evidence of the clinical efficacy of the therapy. These parameters should include a measurement of symptoms and of the use of concomitant medications, which represent the ‘primary outcome’ parameters. Both physician-rated and patient self-rate scores have been implemented in clinical studies. Furthermore, disease-unspecific (generic) and disease-specific questionnaires for evaluating the quality of life are widely used and partially validated as ‘secondary outcome’ parameters. This review provides an overview on the different methods to measure the clinical outcome of SIT and points out the advantages and disadvantages of each method.

Introduction

Allergic rhinitis reveals typical clinical symptoms such as sneezing, rhinorrhea, itching, nasal obstruction and sleep disturbance and, furthermore, impairs the patients' productivity and the well being in general.[1,2••] The treatment of allergic rhinitis comprises allergen elimination, pharmacologic treatment and specific immunotherapy (SIT).[3••,4]

Different bioparameters are used as outcome measures to evaluate the clinical efficacy of SIT. There are two criteria to evaluate the clinical outcome; one criterion is the patient's self-rate score and the other is the physician-rated score.[5•,6]

The clinical monitoring of SIT has the following aims:[2••,3••]

1. to record the patient's condition and the reaction toward SIT, the adjustment of the dosage in particular (safety monitoring);
2. collecting data to continue the therapy (therapy monitoring);
3. collecting trial data for the analysis;
4. evaluating the effectiveness of SIT (quality assessment).

There are primary and secondary endpoint parameters to evaluate the clinical outcome of SIT. Primary endpoints are the severity of symptoms and the need for concomitant medication, which are obtained by diaries.[3••] Secondary endpoints refer to the specific and general quality of life[7-9] and the economic productivity[10] obtained by questionnaires.

In addition to that, there are quite a number of surrogate markers, for example, cytokine analysis, cell-activating markers or proliferation assays, to validate the therapeutic effects of SIT. However, these surrogate markers are only used as additional parameters for the evaluation of SIT or for understanding the immunological mechanisms of SIT and are not used for routine assessments.

This article provides an overview of the currently used and approved methods to assess the clinical outcome of subcutaneous immunotherapy (SCIT) and sublingual immunotherapy (SLIT) in clinical trials and clinical routine. However, there currently is still a lack of national and international guidelines on assessing the therapeutic effect of immunotherapy with clinical markers.

Primary Endpoints in Clinical Trials: Symptoms and Concomitant Medication
The clinical outcome of immunotherapy is assessed by a decrease in the severity of symptoms and a decrease in the need for concomitant medication. Because of their interactions, clinical symptoms and concomitant medication should not be analyzed separately. Therefore, the World Allergy Organization (WAO) claims in a recent statement that the symptom score should always be combined with the medication score.

The patient-related self-assessment regarding the severity of their symptoms is a reliable marker to prove the efficacy of SIT but requires a compliant and motivated patient. For symptom evaluation, diaries are a reliable tool. They should be filled out once or twice daily. Choosing the right diary (or designing a diary) depends highly on the allergist's experience. Following are the practical details for favorable design of patients' self-rated diaries according to the authors' experiences:

1. diaries should be small and handy, to be carried around more easily;
2. water-proofed bindings;
3. only one page for each day;
4. multiple-choice questions and some free lines for free text;
5. huge letters for elderly patients.

For the evaluation of allergic rhinitis, the assessment of the symptoms 'nasal congestion', 'sneezing', 'nasal itching' and 'nasal secretion' are preferably used to establish a nasal symptom score and should be assessed separately. Each symptom has to be documented in a diary, for example, using a four-digit score (Table 1). Therefore, the combined nasal symptom score ranges between 0 and 12 points. According to that, the ocular symptom score should assess 'itching of eyes', 'ocular secretions' and 'redness of eyes' and reaches 0-9 points.

Table 1. Evaluation of Nasal Symptoms Using a Four-digit Score

<table>
<thead>
<tr>
<th>Medscape</th>
<th>0</th>
<th>No symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>Mild symptoms (symptoms with no impact on the patient's life)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Moderate symptoms (regular symptoms with no impact on the patient's daily activity or sleep behavior)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Severe symptoms (symptom-related impairment of the daily activity and sleep behavior)</td>
</tr>
</tbody>
</table>

Adapted from [6,12].

Source: Curr Opin Allergy Clin Immunol © 2009 Lippincott Williams & Wilkins

In recent studies on SLIT with allergen tablets, the Rhinoconjunctivitis Total Symptom Score (RTSS) was used, which included the six most common symptoms of pollinosis (sneezing, rhinorrhea, nasal pruritus, nasal congestion, ocular pruritus and watery eyes) according to the Center for Drug Evaluation and Research guidance. In other studies on SLIT with grass-pollen tablets, a total of 10 types of symptoms were rated in a combined symptom score of nasal symptoms (running, blockage, sneeze, itching), ocular symptoms (gritty/red/itching, watery) and 'pulmonal' symptoms (cough, wheeze, chest tightness/dyspnea, exercise-induced asthma) as the primary endpoint parameter. Each symptom was scored on a scale of 0 to 3 and totaled daily.

Comparing the outcome measures of these studies illustrates that symptom and medication scores currently used in clinical trials as primary outcome parameters are not standardized or evaluated thoroughly yet. For psychometric analysis, the visual analog scale (VAS) is recommended, which is another self-assessment method. VAS has been shown to preferably monitor the severity of symptoms, especially for long-term follow-up in the treatment of perennial allergic rhinitis. For this method, a 10 cm line represents the severity of symptoms from 0, not so troublesome, to 10, unbearably bothersome. The patient marks the severity of his symptoms by putting a vertical line on this VAS (representing the last day, week or month).

There is much evidence that SIT reduces allergic symptoms. However, this therapy abolishes all symptoms only in rare cases. Therefore, clinical studies should guarantee that the patient is supplied with sufficient rescue medication to ease symptoms. Concomitant medications should be ranked according to their ability to reduce symptoms. Following the WAO guidelines, nasal, conjunctival and orally applied antihistaminics should count with 1, inhalative or nasal corticosteroids with 2 and the daily...
application of oral corticosteroids with 3 points. The 'symptom load' of each day is calculated by adding the sum of the symptom score to the medication score to ensure a weighted relationship between symptoms and concomitant medication.\[3\••\]

These scores can be used to determine the total score as the 'area under the curve'. The total score for the verum group can be opposed to the placebo group, and the decrease in the severity of symptoms between these two groups can be verified. Furthermore, patients are asked to retrospectively assess the efficacy of the therapy after pollen season (or after immunotherapy). However, for a clear interpretation of the clinical efficacy of SIT (nonresponder vs. responder), the clinical outcome should be taken into account; a decrease in symptom scores should be high enough to 'really' reduce allergy-related morbidity.\[18\] Comparable measures are either the clinical symptoms, respectively, the use of concomitant medication before treatment (one season for seasonal allergies or a certain period of time for perennial allergies), or the identical parameters of a comparator group ('matched groups').\[2••\]

Referring to this, it is apparent that in some clinical trials, the statistical significance is more emphasized than the clinical relevance of changes. Of course it is important to prove a statistical significant difference between verum and placebo groups, but this does not guarantee the clinical benefit of a therapy in practice under 'real-life' conditions.\[19\] Thus, a 10% reduction in concomitant medication may be statistically significant but does not provide clinical relevance for the patient. Therefore, it is important to establish ranges of symptom or medication effects for a clinical assessment:\[18\]

1. no effect, improvement of less than 30%;
2. little effect, improvement of 30-45%;
3. moderate effect, improvement of 46-60%;
4. strong effect, improvement of more than 60%.

Even though this classification seems to be random, this approach has been found to be a sensitive and reliable tool for clinical assessments.

Most of the clinical trials for SIT present combined symptoms-medication scores but do not consider logistic issues. Contrariwise, in trials assessing the symptom score and medication score separately, this weighted description might distort the interpretation of the results. For example, a decrease in symptoms at a steady use of medication (or vice versa) still results in an improvement in the severity of the disease. However, if symptoms and medication show inconsistent or contradicting results, it might be assumed that there is a lack of a clear clinically relevant improvement of the therapy.

Furthermore, patient self-rated assessments (diaries) for asthma symptoms are feasible. These results might be supported by lung function tests, for example, vital capacity, peak flow or forced expiratory volume in 1 s.\[20\] More than others, the peak flow analysis is a sufficient parameter for asthma monitoring and provides a distinct advantage over the patient self-rated assessments because of its objectivity. Furthermore, inhalative provocation test with methacholine can be used to assess unspecific bronchial hypersensitivity in asthma patients with nearly normal lung function test results despite asthmatic symptoms.\[20\]

Secondary Endpoints in Clinical Trials: Quality of Life

The impact of allergic rhinitis on the patients' health status goes far beyond the nasal symptoms such as itching, sneezing, secretion and obstruction and the ocular symptoms such as itching of the eyes or running eyes alone. Thus, clinical trials on SIT should not only evaluate certain symptoms of the disease but, moreover, also assess the patients' mental and physical condition, the health-related quality of life (HRQoL).\[21\] Clinical trials could clearly demonstrate that the patients' physical fitness, school and work achievement, life satisfaction and the 'daily life' are impaired by allergic rhinitis. Moreover, the sleeping behavior in particular is essential for the HRQoL as sleep impairment results in a noticeable reduction in daily activities.\[22\]

In addition, assessing the HRQoL allows a much more sensitive evaluation of a therapeutic improvement in SIT than just evaluating the symptoms. In a placebo-controlled trial\[23\] on local treatment of allergic rhinitis with budenosid or fluticasone, no significant change in the patients' medication or symptom score could be demonstrated, whereas remarkable differences between these two corticosteroid treatments were found in the 'Rhinconjunctivitis Quality of Life Questionnaire' (RQLQ).\[7\]

There are several health-related questionnaires, such as RQLQ,\[7\] asthma questionnaire or Rhinasthma,\[8,9\] adjacent to generic (health-related) tests, such as Short Form-36 (SF-36) 'Health Status Questionnaire',\[24,25\] the 'Medical outcome study Short Form-20 (SF-20),\[26\] the 'Satisfaction Profile'[27] or the 'Munich Life Dimension List',\[28\] to measure the quality of life\[11,21\] ( Table 2 ).\[29-32\]
Generic questionnaires allow the comparison of patients with different diseases (or healthy individuals) concerning their quality of life. For example, the pharma-economical ‘Work Productivity and Activity Impairment Questionnaire’[^10] measures the degree of the patients’ work impairment. This test assesses:

**Table 2. Quality of Life Questionnaires for Specific Immunotherapy Trials**

<table>
<thead>
<tr>
<th>Medscape</th>
<th>Source: Curr Opin Allergy Clin Immunol © 2009 Lippincott Williams &amp; Wilkins</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generic (not disease related)</strong></td>
<td></td>
</tr>
<tr>
<td>Medical outcome study Short Form-36</td>
<td>Stewart [25]</td>
</tr>
<tr>
<td>Medical outcome study Short Form-20</td>
<td>Carver et al. [26]</td>
</tr>
<tr>
<td><strong>Specific (illness related)</strong></td>
<td></td>
</tr>
<tr>
<td>Asthma Quality of Life Questionnaire</td>
<td>Juniper et al. [29]</td>
</tr>
<tr>
<td>Mini Asthma Quality of Life Questionnaire</td>
<td>Juniper et al. [30]</td>
</tr>
<tr>
<td>‘Munich Life Dimension List’</td>
<td>Kremer et al. [28]</td>
</tr>
<tr>
<td>Asthma Questionnaire 30</td>
<td>Barley et al. [31]</td>
</tr>
<tr>
<td>Asthma Questionnaire 20</td>
<td>Barley et al. [31]</td>
</tr>
<tr>
<td>Rhinoconjunctivitis Quality of Life Questionnaire</td>
<td>Juniper and Guyatt [7]</td>
</tr>
<tr>
<td>Mini Rhinoconjunctivitis Quality of Life Questionnaire</td>
<td>Juniper et al. [32]</td>
</tr>
<tr>
<td>Rhinasthma</td>
<td>Baiardini et al. [8]</td>
</tr>
<tr>
<td>Rhinasthma (adapted German version)</td>
<td>Mosges et al. [9]</td>
</tr>
</tbody>
</table>

Adapted from [3**, 11].

1. frequency of disease-related work absence,
2. impairment of work ability,
3. decrease in productivity,
4. reduction in activity.

The test analysis results in a percentage score, whereas a high percentage score implies a decrease in work productivity. Another commonly used generic questionnaire is the SF-36 ‘Health Status Questionnaire’ that has been widely validated and adapted to many languages.[^24,39]

In contrast to these global, unspecific tests, disease-specific tests were designed in order to assess the HRQoL in certain diseases. In 1991, Juniper and Guyatt[^7] established the RQLQ, a validated disease-specific instrument for evaluating HRQoL on the basis of how symptoms and treatment affect the patient's physical, social and emotional well being.

The RQLQ comprises 28 items in seven domains (activity limitations, emotional function, eye symptoms, nonhayfever symptoms, nasal symptoms, practical problems and sleep problems), all of which are scored by the patients using a seven-point scale ranging from 0 (i.e. not troubled/none of the time) to 6 (i.e. extremely troubled/all the time). Following are the seven main domains that impair the patients’ quality of life with allergic rhinitis assessed by the RQLQ.[^7]

1. practical problems,
2. three individual-stated activities,
3. sleep behavior,
4. emotional condition,
5. nasal symptoms,
6. eye symptoms,
7. nonhayfever symptoms.

The RQLQ sleep domain and the activity domain contain three items each. For sleep, these items are predefined as difficulty getting to sleep, waking up during the night and lack of a good night's sleep. The activity items are freely chosen by each patient from a list of 29 activities. The patient is required to choose the three activities that he deems to be most affected by his

[^10]: Juniper et al.
[^24]: Baiardini et al.
[^3]: Mosges et al.
[^7]: Juniper and Guyatt
rhinoconjunctivitis and then score for the same activities at every follow-up visit.

A recent trial by Ciprandi et al.[5] was able to validate the correlation between the RQLQ and clinical parameters as well as immunological and functional surrogate markers. Thus, this outcome measure represents an important aspect to consider in managing patients with allergic rhinitis. Furthermore, a recent study on SLIT with grass-allergen tablets could demonstrate that the RQLQ scores correlated with both primary parameters (RTSS) as well as secondary parameters (efficacy at pollen peak, combined scores and immunological changes).[34] The disadvantage of these specific questionnaires is that they evaluate allergic rhinitis and allergic asthma as two separate diseases even though they provide the same pathophysiologic basics and often overlap. The 'Rhinasthma' questionnaire, established by Biaiardini et al.[8] for international trials and transferred by Mosges et al.[9] for German trials, enables the investigator to assess the functional, physical and emotional problems of adult patients with allergic rhinitis 'and' bronchial hyperactivity due to the concept of the 'united airways'.

**Surrogate Markers**

Objective measures such as spirometry, paraclinical parameters such as serum-IgE and serum-IgG levels, eosinophil cationic protein (ECP), interferon-γ, interleukin or tryptase levels or provocation tests (nasal, dermal, conjunctival, bronchial) provide additional information but cannot replace the clinical assessment.[3••] Table 3 gives an overview on the paraclinical parameters that are commonly used additionally for the assessment of the benefit of SIT.[3••]

**Table 3. Surrogate Markers and Paraclinical Parameters to Evaluate the Therapeutic Effect of Specific Immunotherapy**

<table>
<thead>
<tr>
<th>Medscape provocation</th>
<th>Immunologic parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organ-related provocation</td>
<td>Total IgE, specific IgE and IgG subclasses</td>
</tr>
<tr>
<td>Skin: titrated skin-prick test</td>
<td>Mucosal IgA</td>
</tr>
<tr>
<td>Nose and eye: provocation test (favorably in titrated doses)</td>
<td>Lymphocyte subgroups</td>
</tr>
<tr>
<td>Lung: spirometric results – FEV1, peak flow variability, metacholine challenge, allergen challenge</td>
<td>Cytokines (e.g. IL-12, IFN-γ, IL-5, IL-10), etc.</td>
</tr>
<tr>
<td>ECP, eosinophil cationic protein; FEV1, forced expiratory volume in 1 s.</td>
<td>Local and systemic inflammatory parameters (e.g. adhesion molecules, leukotrienes in urine, ECP, tryptase)</td>
</tr>
</tbody>
</table>

The nasal provocation test (NPT) is a current standard procedure to reveal the clinical relevance of an allergen for allergic rhinitis. This method is standardized and provides both a high sensitivity and specificity[35] and can either be used as a tool in the diagnosis of allergic rhinitis as well as an objective outcome parameter of SIT.[38] For the latter, the NPT should be performed with increasing threshold concentrations to compare these threshold doses with the results at baseline. In a recently published study on intralymphatic allergen administration in SIT, clinical efficacy was demonstrated by an increased tolerance to nasal provocation with pollen already within 4 months of therapy.[37•]

The provocation test in pollen chambers provides a valuable method to evaluate the therapeutic effect of SIT.[38] So far, this method is not validated thoroughly with the natural exposition during and out of the pollen season.

**Specific Immunotherapy with Hymenoptera Venom**

The evaluation of the therapeutic effect of SIT can be assessed by a controlled sting challenge that is standardized in international guidelines.[39] The sting challenge test under clinical conditions provides a much better prognosis than an accidental 'field' sting.[40••] However, this challenge should not be performed for the diagnosis of sensitization of patients who have not yet received SIT, as, in these patients, severe, life-threatening reactions might occur. Furthermore, negative test results in these patients do not have any prognostic value for an accidental sting in the future.[39]

**Conclusion**

In clinical trials on SIT, several clinical parameters are commonly used to provide evidence of the clinical efficacy of the therapy.
'Primary outcome parameters' include the measurement of symptoms and the use of concomitant medications, whereas quality of life measures or surrogate markers are used as 'secondary outcome' parameters.

References

Clinical Outcome Measures of Specific Immunotherapy (printer-friendly)


Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
•• of outstanding interest

Authors and Disclosures

Oliver Pfaar, a,b Clemens Anders, a,b Ludger Klimek a,b

aCenter for Rhinology and Allergology, Wiesbaden, Germany
bDepartment of Otorhinolaryngology, Head and Neck Surgery, University Hospital Mannheim, Mannheim, Germany

Reprint Address
Dr Oliver Pfaar, Center for Rhinology and Allergology, Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital Mannheim, An den Quellen 10, 65189 Wiesbaden, Germany Tel: +49 611 8804799; E-mail: oliver.pfaar@hno-wiesbaden.de.