Allergol Immunopathol (Madr). 2017;45(S1):17–22

Allergologia et immunopathologia

Sociedad Española de Inmunología Clínica, Alergología y Asma Pediátrica

www.elsevier.es/ai

REVIEW

# Worldwide allergen immunotherapy guidelines: Evidence and experience-based

**D.E.S. Larenas-Linnemann**

*Investigational Unit, Hospital Médica Sur, Mexico City, Mexico*

Available online 3 November 2017

## Introduction

### The history of allergen immunotherapy

Since the first intents of allergen immunotherapy (AIT), now already over a hundred years ago by Leonard Noon, AIT is still the sole disease modifying treatment available for patients with allergic rhinitis and conjunctivitis, allergic asthma, atopic dermatitis and hymenoptera venom allergy. The first publication on AIT in 1911 by Noon from the St. Mary's Hospital in London, described the reduction of sensitivity in nasal challenge testing of hay fever patients, allergic to grass pollen, after repetitive subcutaneous administration of a grass pollen extract.[1] Misfortunately, Noon never could see if the patients really improved their symptoms during the grass-pollen season, as he died of tuberculosis before the start of the season. His colleague, J Freeman, continued his work and subsequently published the reduction in symptoms during the grass pollen season of Noon's AIT patients.[2] Since, AIT has gone through better and worse times, but through the difficult years during the 40–50ies, when AIT was practically banished, it made it to the 60ies when the first controlled trials saw the light, both in the old continent and in the United States of America (USA).[3,4] The first dose–response effect for AIT was shown with Ragweed pollen by Lowell and Franklin.[5,6] During those same years doctor Mary Loveless showed that the administration of wasp and honeybee venom could reduce the frequency and severity of sting reactions in hymenoptera venom allergic patients, as opposed to the poor performance of the whole body extract, typically used in those days.[7,8] The efficacy of AIT in allergic asthma was demonstrated by Johnstone et al. a few years later: multi-allergen AIT reduced the frequency of asthma symptoms and asthma attacks in a dose–response manner.[9]

Since 1986 sublingual AIT showed to be effective as well, and from that moment onward the evidence for this treatment modality has become very solid,[10] demonstrating even long-term effect, be it the evidence here is not as robust as for the direct efficacy.[11]

### The history of guidelines for medical practice

In parallel with AIT, clinical practice guideline-making has also been evolving, but most so only over the past decades, see Box 1. The first guidelines dating from the 70–80ies were mostly consensus documents, narrating recommendations based on the clinical experience of well-recognized experts. By the end of the past millennium evidence based medicine became stronger and Shekelle proposed a system in which the clinical recommendations were directly linked to the level of evidence. In the Shekelle system the quality

*E-mail address:* Marlar1@prodigy.net.mx

https://doi.org/10.1016/j.aller.2017.09.005

**Box 1**

Evolution of the methods to develop guidelines over the past decades

- -1990 'Expert opinion'
- 1990 -2005 EBM = evidence based medicine (Shekelle*)
  - Evidence categories-IIV
    - Ia Evidence from metaanalysis of randomized, controlled clinical trials (RCTs**)
    - Ib Evidence from at least one RCT**
    - IIa Evidence from at least one controlled trial–no randomization-.
    - IIb Evidence from at least a quasi-experimental study
    - III Evidence from descriptive, non-experimental studies, e.g. comparative trials
    - IV Evidence from reports of expert committees and/or clinical experience from respected authorities
  - Strength of recommendation:
    - Directly based on category I evidence
    - Directly based on category II evidence or extrapolated from category I evidence
    - Directly based on category III evidence or extrapolated from category II evidence.
    - Directly based on category III evidence or extrapolated from category IV evidence.
    - LB Laboratory-Based on reasoning from first principles (bench research or animal studies)
    - NR Not rated
- 2005 - GRADE
  - Quality of scientific evidence + other factors

of the evidence was directly related to the study design, with metanalyses having the highest category of evidence (1a), directly followed by randomized clinical trials (category 1b). As such, in the Shekelle system it is possible to respond to a certain clinical question with a level A recommendation, based on only one clinical trial, in which active and control patients are openly randomized, without the need for a placebo group. In this same line of acting a doubleblind, placebo controlled clinical trial, even though completely underpowered, can be able to generate a level A recommendation in the Shekelle system.

That's why a group of methodologists from the McMaster University in Ontario, Canada, developed in 2005 the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system, which has been the leading system of guideline development in all areas of the medical discipline since. In GRADE the recommendations are the core part of the system, and the evidence is only one pillar that sustains a recommendation. Apart from the evidence, the safety of the treatment, the cost and the patients' preference are the other three pillars on which the recommendations are built. Moreover, the quality evaluation of the evidence is more complete, taking into account other factors, apart from the study design, that can help to enhance or reduce quality of a study, see Fig. 1. As a consequence, evaluating with GRADE not all double-blind placebo controlled trials are of high quality, nor all observational studies of low quality, and all can contribute to the evidence needed in a guideline.

Even though the GRADE system also has its flaws and is still in continuous improvement, since its creation in 2005, it has been considered the most reliable guideline development system.

The evidence can be assigned from 1 to 4 bullets, rating it from very low to high quality evidence. The study design is only the starting point – see left part – from where quality points can be added or subtracted according to further details in study design, elaboration of the results and publication.

## Assessing the quality of medical guidelines: the AGREE II tool

With the appearance of more and more guidelines dealing with a variety of medical issues, from global ones, to regional and national ones lately two further problems with guidelines are becoming apparent: 1. The quality, 2. The poor applicability. To help the medical community differentiate between the quality of the different guidelines, again methodologists of the McMaster University developed a tool to grade guidelines quality: AGREE, in 2010 improved into AGREE II.

AGREE II consists of six domains, that evaluate:

- Domain 1. Scope and Purpose
- Domain 2. Stakeholder Involvement
- Domain 3. Rigor of Development*
- Domain 4. Clarity of Presentation
- Domain 5. Applicability
- Domain 6. Editorial Independence

*(process to gather and synthesize the evidence, the methods to formulate the recommendations and to update them)

Each domain consisting of several questions that investigators should pose to the guideline; for example 'The guideline development group includes individuals from all relevant professional groups'. Each item is evaluated from 1 (Strongly disagree) to 7 (Strongly agree).

One of the strong parts of AGREE II is the emphasis it puts on the enhancement of guideline dissemination and applicability. Already in the second domain it stresses the importance of a broad stakeholder involvement: right from the early phases of the development of a guideline not only the specialists, but also primary care doctors, pharmacists, nurses and industry partners and legislators should be involved. This is re-enforced with the fifth domain: applicability, that asks for a separate chapter in the guideline