Method

In order to standardize the method for finding and classifying evidence and the recommendations in this guide, all the editors were given training in methodological aspects by collaborators from the Cochrane Latin American Center. In order to identify publications, the usual procedure for preparing clinical practice guidelines was followed [1] and the reference lists of the main international clinical practice guidelines were reviewed [2-4] in order to identify the main systematic reviews and clinical trials. These guidelines were found in specialized databases (National Guideline Clearinghouse, National Library of Guidelines) and in the TRIP medical literature meta-search engine database. The databases of the Centre for Reviews and Dissemination (DARE and HTA database) and The Cochrane Library were consulted to identify systematic reviews and evaluations of additional technologies. To complete the search the data was updated to include any systematic reviews and relevant studies in the major electronic databases of original studies (MEDLINE, CENTRAL and EMBASE) that were published after the search dates.

To assess the quality of the evidence, an alphabetical classification, which graded the quality of the information into four categories (A, B C, D), was employed (Table 0.1). This classification reflects the level of confidence in the results that were obtained in the studies available [2]. Category A corresponds to a high level of quality and category D to a very low level of quality. The confidence in the results, in the case of category A, is high and it is improbable that subsequent studies could modify their findings. In contrast, in the case of the lower categories, such as C or D, confidence is low or very low, and it is very likely that later studies could alter the results or even the rationale underlying them.

However, as the editors of this guide, we understand that although this classification is very useful for categorizing evidence designed to assess the therapeutic efficacy of drugs or other interventions, it tends to underrate other equally important studies, such as trials on the diagnostic efficacy of certain tests or epidemiological data. This is why in this guide much of the evidence available for evaluating important studies on the suitability of certain diagnostic tests has been graded with a C.

Given the recent appearance of new approaches to classifying the quality of evidence based on other aspects in addition to study design [5], future updates of GEMA will try to reflect these changes in the way it develops its recommendations. This time certain conceptual aspects and basic features of the GRADE system (http:// www.gradeworkinggroup. org/) have been included, although this system has not been applied in any strict sense [5].

Once the quality (confidence in the results) of the studies at our disposal had been classified, the strength of our recommendations had to be classified, meaning our confidence that compliance with a particular recommendation would lead to more benefits than risks. To establish the strength of recommendations, the quality of the information (based on the above mentioned classification), the riskbenefit ratio of interventions, the costs and the values and preferences of patients were considered. Then, the recommendations were classified into two types: strong and weak recommendations (in favour of or against). Strong recommendations (R1 recommendations) are those the group that drew up the guidelines are convinced will be associated with more benefits than risks. For this type of recommendation the text uses expressions such as “it is recommended” or “we ought to”. Weak recommendations (R2 recommendations) are those for which there is uncertainty as to whether their application would entail more benefits than risks and the language they use includes expressions such as “it might be considered” or “it could be regarded as”.

Table 0.1. Classification of the quality of the evidence obtained from searches

Evidence Category	
A	SR of RCT with or without MA and RCT with a low risk of bias. The evidence is based on a substantial number of well designed studies with consistent results.
B	SR of RCT with or without MA and RCT with a moderate risk of bias. The evidence is based on a limited number of studies and/or inconsistent results.
C	The evidence is from non-randomized, observational or uncontrolled studies.
D	Clinical experience or scientific literature hat cannot be included in category C.

RCT: randomised clinical trials; MA: meta-analysis; SR: systematic reviews (modified from GINA 2006) [2].