

Appendix 2

Delphi Process

Pertinent literature reviews were sent to all of the committee members. Subsequently, a rating survey, based upon the literature review, was developed by a subgroup of the committee members, the committee co-chairmen and the methodologist. This questionnaire consisted of 79 rating statements, 39 of which were related specifically to a clinical classification scheme for early stages of AMD and the remainder to utilizing additional techniques beyond a simple ophthalmoscope as part of a classification scheme. This was posted on a designated secure website at which the committee members could rate their level of agreement with each statement anonymously and independently. The items were to be rated 1-9 with 1 being “strong disagreement” and 9 representing “strong agreement”, and 5 being “neither agree nor disagree”. The 9-point scale was identified as having 3 broad zones: disagreement (1 to 3); neutral (4 to 6); and agreement (7-9).^{11,12}

The extent of consistency, or consensus, in ratings by the committee members regarding their degree of agreement with the statements formed the basis of committee discussion at the first committee meeting. Median committee ratings, as well as the distribution of ratings for each statement, were calculated, and consensus and non-consensus were determined by applying a statistical definition with a binomial distribution.^{11,12} Consensus was defined as no more than 20% of individual ratings outside the 3-point bracket containing the median, and it could include consistency to agree or not to agree with a given statement. Non-consensus was defined as \geq 33% of committee responses in polar ranges (1-3 or 7-9). Equivocal included all others. “Agreement” for inclusion as criteria included a committee median of 7-9 without non-

consensus, whereas “disagreement” included a committee median of 1-3 without non-consensus. “Uncertain” included a committee median of 4-6 or any median with non-consensus.

The results were tabulated to identify the median levels of grading of each of the statements and to establish areas of consensus, non-consensus, and “equivocal” positions. The results were returned to each responding committee member. These contained the anonymous responses of all, a comparison of the vote of the respective member to that of the group, and the level of consensus among the members of the committee. A first meeting of the group occurred in August 2011, with 29 individuals attending. At the initial session, the committee members discussed the outcomes of the initial ratings with an emphasis on statements that were rated as non-consensus or intermediate. The Delphi method focuses on the stability of group opinions rather than on the opinions of individuals in an effort to assess the degree of consistency of judgment among committee participants^{11,12}, and the goal of the discussion was not to force committee members into greater consensus but to give them the opportunity to consider the differing opinions of members.