ACR Appropriateness Criteria® Rating Process

INTRODUCTION

The ACR Appropriateness Criteria® Rating Process is based on the RAND/UCLA Appropriateness Method User's Manual [1]. It is a relatively simple process but complex in its implementation. The complexity arises from the desire that those rating appropriateness do so from a "level playing field" to the extent possible. This way the only differences in rating will be based on the understanding of the evidence from the medical literature, or when this evidence is lacking or contradictory, the evidence from the medical experts' knowledge, expertise, and experience.

PROCESS SUMMARY

The purpose of the rating process is to systematically and transparently determine the appropriateness of performing an imaging procedure or treatment for a specific clinical scenario while focusing on the available evidence and mitigating any unintended influence or persuasion among panel members.

In addition to evaluating appropriateness, the purpose of the appropriateness rating process is to determine whether the experts disagree regarding the final appropriateness rating category. This is achieved using a modified Delphi method, based on the RAND/UCLA appropriateness method [1], used to formulate the recommendations through consensus regarding interpretation of the available evidence. The transparency of the AC methodology exposes biases in appropriateness ratings while allowing experts with different clinical expertise to indicate their interpretation of the evidence independently.

The ACR adopted the definition of appropriateness mentioned in the RAND/UCLA Appropriateness Method User's Manual [1]:

> *...the expected health benefit (eg, increased life expectancy, relief of pain, reduction in anxiety, improved functional capacity) exceeds the expected negative consequences (eg, mortality, morbidity, anxiety, pain, time lost from work) by a sufficiently wide margin that the procedure is worth doing, exclusive of cost"* (Brook et al., 1986; Park et al., 1986).

Panel members use the evidence summary to rate recommendations based on the benefits and potential risks of performing an imaging procedure or treatment for the specific clinical scenario. In assessing the harms and benefits, panel members focus on how well the imaging procedure or treatment can provide clinical information to move the patient along the care pathway to the best outcomes. This should be based on similar patient population as defined by the variants. Panel members will often need to apply their expertise to evaluate the benefit of the procedure or intervention for the variant based on evidence for populations and clinical scenarios that may have different characteristics.

The appropriateness category names were modified from the original categories in the RAND/UCLA Appropriateness Method User's Manual [1] The appropriateness rating range is an ordinal scale of integers from 1 to 9 grouped into three categories; "Usually not appropriate" (1, 2, or 3), "May be appropriate" (4, 5, or 6), and "Usually appropriate" (7, 8, or 9) to account for those instances where specific patient or other characteristics presented to the ordering provider who using clinical judgment, may modify the final decision. At the ends of the rating scale are the "usually not appropriate" category, which describes when the potential risks of doing the procedure outweigh the benefits and the "usually appropriate" category, which describes when the benefits of doing a procedure outweigh the potential risks. The middle "may be appropriate" category is applicable when the risks and benefits are equivocal or unclear, the dispersion of the individual ratings from the group median rating is too large (ie, disagreement), there are special circumstances or subpopulations that are embedded in the variant which confound the benefits-risks assessment or the evidence is contradictory or unclear. The evidence for a recommendation does not directly

ACR Appropriateness Criteria® Rating Process

affect the determination of panel disagreement or agreement; the rating process is a method to objectively determine if group consensus has been achieved.

Panelists rate each of the procedures in up to two-rounds. The ratings are anonymized to allow experts with different clinical expertise to indicate their interpretation of the evidence. If there is no disagreement, the panel's appropriateness recommendation is determined by the median of the panelists' ratings. If there is wide variation of the individual ratings from the group median rating (disagreement) after the first round, a conference call takes place to determine if the clinical scenarios are clear or misunderstood and if the evidence is being considered by all. A second rating round takes place sometime after the call has concluded. If there is still disagreement after two rounds, the rating is categorized as "May be appropriate" and "5" is the assigned rating.