

**Appendix 3: Appropriateness Rating Process**

The rating process used in the rating rounds is a modified Delphi method based on the methodology described in the RAND/UCLA Appropriateness Method User Manual<sup>i</sup> where “the expected health benefit (e.g. increased life expectancy, relief of pain, reduction in anxiety, improved functional capacity) exceeds the expected negative consequences (e.g. mortality, morbidity, anxiety, pain, time lost from work) by a sufficiently wide margin that the procedure is worth doing, exclusive of cost”.

Rating scores modeled on the American College of Radiology Appropriateness Criteria are in an integer ordinal scale from 1 to 9 in three categories. Ratings 1, 2, or 3 are in the “Usually not appropriate” category, where the harms of undergoing a diagnostic procedure outweigh its benefits and 7, 8, or 9 are in the category “Usually appropriate” where the benefits outweigh the risks. The middle category is called “May be appropriate” and is represented by 4, 5, or 6 on the scale. The middle category describes when the risks and benefits are equivocal or unclear, the dispersion of the individual ratings from the committee rating is too large, the evidence is contradictory or unclear, or there are special circumstances or subpopulations which could influence the risks or benefits that are embedded in the variant.

Appropriateness Category Name	Rating Score	Appropriateness Category Definition
Usually Appropriate	7, 8, or 9	The diagnostic procedure or intervention is indicated at a favorable risk-benefit ratio for patients in the specified clinical scenarios.
May Be Appropriate	4, 5, or 6	The diagnostic procedure or intervention may be indicated as an alternative to more favorable risk-benefit ratio in the specified clinical scenarios, or the risk-benefit ratio for patients is equivocal.
May Be Appropriate (with Disagreement)	5	Majority of the individual committee member’s ratings deviate away from the committee’s median score. “May be appropriate” remains the category but with a rating score of 5 to provide transparency regarding the committee’s recommendation.
Usually Not Appropriate	1, 2, or 3	The diagnostic procedure or intervention is unlikely to be indicated in the specified clinical scenarios, or the risk-benefit ratio for patients is likely to be unfavorable.

<sup>i</sup> Fitch K. The RAND/UCLA appropriateness method user's manual. Santa Monica: RAND Corporation; 2001.