

## ACR Appropriateness Criteria® Rating Round Information

### Overview

The purpose of the rating rounds is to systematically and transparently determine the panels' recommendations while mitigating any undue influence of one or more panel members on another individual panel members' interpretation of the evidence. The panel member's rating is determined by reviewing the evidence presented in the Summary of Literature Review and assessing the risks or harms of performing the procedure or treatment balanced with the benefits of performing the procedure or treatment. The individual panel member ratings are used to calculate the median rating, which determines the panel's rating. The assessment of the amount of deviation of individual ratings from the panel rating determines whether there is disagreement among the panel about the rating.

The process used in the rating rounds is a modified Delphi method based on the methodology described in the RAND/UCLA Appropriateness Method User Manual ([Fitch, 2001](#)).

The appropriateness is rated on an ordinal scale that uses integers from 1 to 9 grouped into three categories: 1, 2, or 3 are in the category "Usually not appropriate", where the harms of doing the procedure or treatment outweigh the benefits; and 7, 8, or 9 are in the category "Usually appropriate" where the benefits of doing a procedure or treatment outweigh the harms or risks. The middle category is called "May be appropriate" and is represented by 4, 5, or 6 on the scale ([Table 2](#)). The middle category describes when the risks and benefits are equivocal or unclear, the dispersion of the individual ratings from the panel 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.

### General Information

- Unless otherwise stated, the panel assesses the appropriateness of the *initial* imaging procedure or treatment for the variant. A single variant table should not attempt to provide a recommendation for the initial imaging procedure or treatment and the subsequent imaging procedure or treatment. The variant table for the subsequent imaging procedure or treatment must be separate from the initial imaging recommendations and must clearly state any prior imaging tests or treatment results that are part of the variant.
- All panel members (including the society representatives) are expected to participate in all of the rating rounds. If a panel member feels he or she does not have the relevant expertise to rate a particular topic, the panel member must notify the ACR staff as soon as possible but no later than the beginning of the rating rounds. The name of the panel member (or the society representatives) will be removed from the authorship for any topic in which she/he does not participate.
- Panel members must review all information included on the evidence table (ET) and draft narrative to assign the appropriateness ratings for each procedure or treatment. The ratings are based on the assessment of the risks and benefits of performing the procedure or treatment for the specified clinical scenario using the available evidence. Expert opinion is required to interpret the evidence and may determine the rating when the evidence is incomplete, contradictory, or nonexistent.
- For existing topics being revised, the panel rates only those recommendations that have new and significant evidence, added or modified procedures, techniques or treatments, or changes in practice behaviors.
- Two rating rounds are conducted typically. Those recommendations that have no disagreement after the first rating round are not rated in the second round unless two-thirds of the panel decides the appropriateness category does not reflect the evidence. A conference call is held after the first rating round to discuss the evidence, the rating trends, and the clarity of the document. Specific discussions to finalize the recommendations on the call are not permitted.
- The panel members access the following materials for rating rounds. See [Table 1](#) below.

**Table 1. Materials used for rating rounds**

Process	Panel member access through GRAVITAS
Rating Round 1 (or first time rating a recommendation)	<ul style="list-style-type: none"> <li>▪ Draft document</li> <li>▪ Evidence Table (ET)</li> <li>▪ Rating Round 1 form</li> </ul>
Conference Call between Rating Round 1 and 2	<ul style="list-style-type: none"> <li>▪ Revised draft document</li> <li>▪ Evidence Table (ET)</li> <li>▪ Results from Rating Round 1 including comments</li> </ul>
Rating Round 2 (or second time rating a recommendation)	<ul style="list-style-type: none"> <li>▪ Most current draft document</li> <li>▪ Evidence Table (ET)</li> <li>▪ Rating Round 2 form including revised comments and notes from call</li> <li>▪ Results from Rating Round 1</li> </ul>
After Rating Round 2 (or after the final rating round)	<ul style="list-style-type: none"> <li>▪ Final draft document</li> <li>▪ Final Rating Round Results</li> </ul>

### Ratings Assumptions

The ACR adopted the definition of appropriateness mentioned in the RAND/UCLA Appropriateness Method User’s Manual ([Fitch, 2001](#)) where “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).

Some risks or benefits may vary due to factors that are unique to a specific patient and provider situation. For example, MRI may be contraindicated in some patients (eg, ferrous metal inside the body). Individual patient characteristics are important in the final selection of the right procedure or treatment for any specific patient. Therefore, the provider and the patient may decide to use a modality that is best for the patient’s situation even when the medical evidence and expert opinion strongly support the use of a different modality for that described clinical scenario.

The following list explains guidelines and assumptions that are to be used by the panel members when rating appropriateness. A more detailed explanation may be found in [Appendix 1](#).

When assigning appropriateness ratings to procedures or treatments, panelists should assume that:

- The primary assessment for diagnostic imaging is the diagnostic utility, diagnostic accuracy, test performance, etc. of performing the imaging procedure or treatment for the described clinical scenario.
- The patients (as defined in the topic and variant) do not have contraindications for any of the procedures or treatments listed in the variant table.
- All procedures or treatments in the variant table are available and accessible.
- All procedures or treatments in the variant table are performed and interpreted by an expert.
- The direct or indirect costs of a procedure or treatment are not considered as a benefit or harm when determining appropriateness.
- The Relative Radiation Level (RRL), radiation exposure, radiation dose or the amount of radiation produced from performing a procedure in the variant table are not considered when determining appropriateness as the calculations can vary widely for a specific individual. The overall benefits of the procedure should be assumed to outweigh any theoretical risk due to potential exposure to ionizing radiation when determining the appropriateness category. Relative radiation levels **may** be considered when assessing risks and benefits **only when** other medical imaging examinations or procedures have nearly equivalent diagnostic accuracy or test performance.

### Determining the Panel’s Recommendation

- Ratings represent an individual’s assessment of the risks and benefits of performing a specific procedure for a specific clinical scenario on an ordinal scale. The recommendation is the appropriateness category (ie, “Usually appropriate”, “May be appropriate”, or “Usually not appropriate”).

- The appropriateness category for a procedure and clinical scenario is determined by the panel’s median rating without disagreement (see below for definition of disagreement). The panel’s median rating is calculated after each rating round. If there is disagreement after the second rating round, the rating category is “May be appropriate (Disagreement)” with a rating of “5” so users understand the group disagreed on the final recommendation. The actual panel median rating is documented to provide additional context.
- Disagreement is defined as excessive dispersion of the individual ratings from the group (in this case, an AC panel) median as determined by comparison of the Interpercentile Range (IPR) and the Interpercentile Range Adjusted for Symmetry (IPRAS). In those instances when the IPR is greater than the IPRAS, there is disagreement. For a complete discussion, please refer to chapter 8 of the RAND/UCLA Appropriateness Method User Manual ([Fitch, 2001](#)).
- Once the final recommendations have been determined, the panel reviews the document. If two thirds of the panel feel a final recommendation is wrong (eg, does not accurately reflect the evidence, may negatively impact patient health, has unintended consequences that may harm health care, etc.) and the process must be started again from the beginning.

**Table 2. Appropriateness Category Names and Definitions**

Appropriateness Category Name	Appropriateness Rating	Appropriateness Category Definition
Usually Appropriate	7, 8, or 9	The imaging procedure or treatment is indicated in the specified clinical scenarios at a favorable risk-benefit ratio for patients.
May Be Appropriate	4, 5, or 6	The imaging procedure or treatment may be indicated in the specified clinical scenarios as an alternative to imaging procedures or treatments with a more favorable risk-benefit ratio, or the risk-benefit ratio for patients is equivocal.
May Be Appropriate (Disagreement)	5	The individual ratings are too dispersed from the panel median. The different label provides transparency regarding the panel’s recommendation. “May be appropriate” is the rating category and a rating of 5 is assigned.
Usually Not Appropriate	1, 2, or 3	The imaging procedure or treatment is unlikely to be indicated in the specified clinical scenarios, or the risk-benefit ratio for patients is likely to be unfavorable.

## APPENDIX 1

### Rating Assumptions

In assessing the harms and benefits, panel members focus on the diagnostic utility, diagnostic accuracy, test performance, etc. of a specific procedure for a particular clinical condition. While finding good quality peer-reviewed journal articles that address the performance and interpretation of a distinct procedure for the diagnosis of a well-defined clinical scenario would be ideal, it is more likely that the articles will be less specific. Expert opinion is required to apply the available evidence to each topic-variant combination.

Each panel member is tasked to determine the appropriateness of the procedure for the typical case in an “ideal” world. By ideal world, we mean that some variables that an individual patient may consider with their physician are assumed to be the same for all patients. These variables (such as costs, availability, accessibility, etc.) may have a value component which is difficult to quantify or assess across the population of those who are likely to have the clinical condition. For example, an individual may live 5 minutes away from where a procedure is performed but feels that is too far. While another person who lives 50 miles away from the procedure may decide the distance to have the procedure is not an issue. Other variables (such as relative radiation levels and contraindications) are more relevant to specific groups within the population. In some instances, a new variant may be created to describe the specific group within the population.

Clarifications about these assumptions will facilitate all panelists to consider appropriateness in a consistent manner. It also mitigates over- or under-estimating the impact of the variables especially as they are combined to a real world case. For example, if a child needed a procedure and the choice was between ultrasonography and computed tomography, some may assume that the ultrasound should be selected because there is no radiation for the child. However, if the CT is superior in the diagnostic accuracy, could be performed with low radiation dose, and it was known that the ultrasound would require expertise beyond what the facility could offer; perhaps the provider and patient would select CT. In the real world, the assessment by the physician will include consideration of these variables in consultation with the patient to designate the best diagnostic procedure for the specific situation.

#### 1. *Relative radiation levels*

When rating appropriateness, **the panelist does not consider relative radiation level of the procedure or treatment.** Relative radiation levels may be considered when assessing risks and benefits only when other medical imaging examinations or procedures have nearly equivalent diagnostic accuracy or test performance.

#### **Explanation/Clarification**

The calculation of radiation dose is complex and outcomes of any radiation dose are specific to each individual. The overall benefits of the procedure should be assumed to outweigh any risk due to potential exposure to ionizing radiation when determining the appropriateness category. The document must clearly state the imaging test performance metrics as compared with the study using ionizing radiation in order to use relative radiation levels to assess appropriateness. In those instances where there are no test performance metrics, relative radiation levels cannot be considered in the appropriateness assessment.

#### 2. *Procedures that depend on the expertise of the practitioner.*

When rating appropriateness, **the panelist assumes that all procedures or treatments are performed by an expert.**

#### **Explanation/Clarification**

Some procedures may be appropriate but the benefits may depend on how well the operator or practitioner performs the procedures or interprets the results. Additional explanation may not be discussed in the narrative text that the results are operator dependent or requires special training or expertise for performance, supervision, or interpretation.

#### 3. *Availability of the procedure*

When rating appropriateness, **the panelist assumes that all procedures or treatments listed on the variant tables are available.**

### **Explanation/Clarification**

The rating should only be based on the evidence of the appropriateness of performing the procedure. The lack of the availability or accessibility of one procedure should not enhance (or diminish) the appropriateness of another procedure. If the procedure is not generally available, additional explanation may not be discussed in the narrative text that suggests which procedure may be optimal in the event that another procedure (which may be more appropriate) is likely not to be available.

#### *4. Procedures that are contraindicated*

When rating appropriateness, **the panelist assumes that the patients, as defined by the topic and variants, do not have contraindications for any of the procedures or treatments.**

### **Explanation/Clarification**

The rating should only be based on the evidence of the appropriateness of performing the procedure for the described topic and variant. Contraindications should not enhance or diminish the appropriateness of any procedure within a variant. If the procedure in a variant is generally contraindicated, a detailed explanation may not be discussed in the narrative text.

#### *5. Cost of the procedure*

When rating appropriateness, **the panelist does not consider the cost of the procedure or treatment.**

### **Explanation/Clarification**

Cost or insurance coverage issues may not be discussed in the narrative text.