AUC: Evidentiary Review Process

A Systematic Literature Review

Our Appropriate Use Criteria have been in general use for over 10 years[1]. The following process was followed for the original creation of the AUC and is used in their continuing maintenance. All AC are on a yearly review cycle. In addition, ad hoc reviews might be done as the result of a request by one of the users, or in response to major publications coming that have an impact upon AUC.

At the beginning of the process, the multidisciplinary AUC team that is responsible for developing and modifying the AUC designates a review committee and a committee chair, usually a subspecialist radiologist.

At MGH the review teams are comprised of multiple panels of subject matter exerts. Each panel includes at least one subspecialist radiologist (e.g. neuro-radiologist, pediatric radiologist, etc.), as well as one clinical specialist (e.g. neurologist/neurosurgeon or pediatrician) and one primary care physician. Other expertise is called in as needed. For example, oncologists may be required to review AUC across multiple organ-systems. At any given time, there may be multiple such committees at work, depending upon the topics requiring review, each reporting to the AUC team.

MGH staff prepares a literature search for the review committee identifies keywords and other search parameters to be used in the literature search based on the topic to be reviewed. A comprehensive literature search using standard indices such as PubMed and MEDLINE to identify relevant peer reviewed medical literature.

- The two general classes of keywords used in the search are those related to the condition (e.g. back pain, fever) and those that describe the diagnostic or therapeutic intervention of interest (e.g. CT, aspiration).
- The search terms and parameters are combined to produce the most relevant articles to address the AC topic.
- For topic updates, the search is restricted to one year prior to when the last update was performed, unless otherwise specified by the author.
- For new topics, the search is restricted to the last ten years, unless otherwise specified by the author.
- The search results are limited to original research articles that have abstracts and human subjects and to meta-analyses.
- Review articles and case reports are excluded from the literature search results.

---

[1] Rosenthal DI, Weilburg JB, Schultz T, et al. Radiology order entry with decision support: initial clinical experience. *J Am Coll Radiol* 2006;3:799–806

The literature search is provided to the review committee, along with a summary statement that includes the time period covered by the search, the database(s) used for the search, a summary of the terms used, the number of articles identified in the search, the number of articles from the search included in the AC topic, and a summary of inclusion and exclusion criteria. This committee supplements or edits the search as it sees fit and may also request that staff conduct additional searches. The literature search is used to draft or revise the AC topic and appropriate citations are embedded in the document.

Methodology

Evidence is weighted using the Grades for Recommendations, Assessment, Development and Evaluation or GRADE methodology[2].

Description of the Methods Used to Analyze the Evidence

The review committee assesses the literature then drafts or revises the narrative summarizing the evidence found in the literature. MGH staff drafts an evidence table based on the analysis of the selected literature. These tables rate the study quality for each article included in the narrative.

The expert panel reviews the narrative, evidence table and the supporting literature for each of the topic-variant combinations and assigns an appropriateness rating for each procedure listed in the variant table(s). Each individual panel member assigns a rating based on his or her interpretation of the available evidence.

The Evidence Table (ET) summarizes the sources cited in the AC topic narrative. Additionally, the ET quantifies a source's quality based on the number of study quality elements described in that source.

Description of the ET components

The ET includes five components extracted from a source. These are the reference citation information, study type, number of patients or events, study objective(s), and study result(s). The study type designates the source's purpose and design. The purpose of diagnostic studies is to diagnose or assess patients by utilizing diagnostic tools while therapeutic studies assess the use of treatments and interventions in treating patients. Furthermore, diagnostic and therapeutic studies have different study quality elements that help assess the amount of bias that may be introduced, which may affect the results and conclusions of the study.

Four study design categories are considered in developing the AUC: experimental, observational, review/other, and meta-analysis. These broadly defined study designs contribute to the assessment of study quality. Well-designed and well-executed experimental studies typically are better at controlling biases and determining causality where other study types, like observational studies, may determine only when there is a relationship between events and outcomes.

Additional information for categorizing these study types is described below.

[2] GRADE Working Group 2004, Schünemann 2006b, Guyatt 2008a, Guyatt 2008b

Study Type Categories

Experimental studies create differences in the explanatory (independent) variable under controlled conditions and examine any resulting changes in the response (dependent) variable. These studies include methodologies that reduce the potential for bias, for example, randomization, blinding. An example is the randomized controlled trial.

The primary characteristics of experimental design are 1) true experiments have control and manipulation, 2) specifies an experimental group and control group, and 3) test cause and measure effect.

For observational studies, investigators observe subjects and measure variables of interest (independent variables) without assigning treatments, interventions, or outcomes to the subjects. The treatment, intervention, or outcome that each subject receives is determined beyond the control of the investigator.

The primary characteristics of observational design are 1) investigator observes variables, 2) specifies cohorts (groups with similar characteristics of interest) or a case (groups with the variable of interest) and control (groups without the variable of interest) group, and 3) tests association between variables but not causality.

Reviews or other studies include case reports, systematic literature reviews, clinical practice guidelines, book chapters, etc. These articles may or may not be studies but include published literature that examines or reviews other studies, data, surveys, opinions, etc. and summarizes results or concludes outcomes. Other types of articles in this category may be studies that have descriptive statistics only that do not provide a result or outcome to the study, such as incidence or prevalence studies that only describe population or disease trends or patterns. Meta-analyses aggregate information in order to achieve a higher statistical power for the measure of interest, as opposed to a less precise measure derived from a single study. To accomplish this, well=designed meta-analyses make use of rigorous assessment of the study design and quality that they aggregate. Other methods that do not create pooled samples using statistical methods such as systematic literature reviews and clinical practice guidelines are not included in the definition of meta-analysis.

The primary characteristics of meta-analysis design are 1) systematic and unbiased review of literature 2) Exclusion of all articles that do not meet the requirements for quality of evidence 3) a pooled analysis of data that meets the quality standard.

Strength of Evidence

Moving forward, MGH will list the strength of evidence for each recommendation on the updated variant table. The references associated with the recommendation will also be displayed.

The strength of evidence assessment is based on the GRADE methodology[3]. Based on an algorithm, the number, quality, directness, and consistency are evaluated to categorize the strength of evidence. The categories are Strong, Moderate, Limited, Expert Consensus, and

[3] GRADE Working Group 2004, Schünemann 2006b, Guyatt 2008a, Guyatt 2008b.

Expert Opinion. The strong category indicates that there are two or more high quality studies that have consistent results and are directly related to the recommendation. The moderate rating indicates good quality studies but there may be insufficient number, quality, directness, or consistency among the studies. Limited strength of evidence indicates that while there is some evidence in the literature, it relies heavily on experts to interpret the meaning and application to practice. Expert consensus is when there are no studies associated with the recommendation but the panel is able to reach consensus for the recommendation. Expert opinion is assigned to a recommendation when there is no consensus among the panel, regardless of the strength of evidence. The interpretation is that even with strong evidence, if the experts cannot agree on a recommendation, the evidence is flawed.