ORIGINAL ARTICLE

The Evidence Value Matrix for Diagnostic Imaging

David Seidel, PhD[a], Richard A. Frank, MD, PhD[b], Sebastian Schmidt, MD[c]

Abstract

Evidence and value are independent factors that together affect the adoption of diagnostic imaging. For example, noncoverage decisions by reimbursement authorities can be justified by a lack of evidence and/or value. To create transparency and a common understanding among various stakeholders, we have proposed a two-dimensional matrix that allows classification of imaging devices into three distinct categories based on the available evidence and value: "question marks" (low value demonstrated in studies of any evidence level), "candidates" (high value demonstrated in retrospective case-control studies and smaller case series), and "stars" (high value demonstrated in large prospective cohort studies or, preferably, randomized controlled trials). We use several examples to illustrate the application of our matrix. A major benefit of the matrix includes the development of specific strategies for evidence and value generation. High-evidence/low-value studies are expensive and unlikely to convince decision makers, given the uncertainty of the impact on patient management and outcomes. Developing question marks into candidates first and then into stars will often be quicker and less expensive ("success sequence"). Only this more sophisticated and objective approach can justify the additional funding necessary to generate the evidence base to inform reimbursement by payers and adoption by providers.

Key Words: Diagnostic imaging, evidence, value

J Am Coll Radiol 2016;■:■-■. Copyright © 2016 American College of Radiology

Health care systems around the world are under acute financial pressure. Governments and payers explore every opportunity to reduce costs. For example, as a result of the Deficit Reduction Act in the United States, overall costs of diagnostic imaging to Medicare Part B were cut by 21%, from $11.9 billion (US dollars) in 2006 to $9.5 billion in 2010 [1]. At the same time, the growth in use of imaging for Medicare beneficiaries slowed down to 1-3% from previously >6% annually, with a meaningful fraction of the decline involving imaging procedures with "unproven value" [2].

Assessing the value of diagnostic imaging has long been a concern. Though there are many factors that influence the adoption of imaging technologies by various stakeholders, "the ultimate basis of acceptance and use is strong clinical evidence supporting the diagnostic value" [3]. The recent decision by CMS to cover lung cancer screening with low-dose CT (LDCT) in high-risk individuals [4] exemplifies this very well: in a multicenter randomized controlled trial (RCT) of more than 53,000 heavy smokers, a 20% reduction in lung cancer mortality was demonstrated in subjects who underwent three annual CT scans compared with chest radiography after a median follow-up duration of 6.5 years [5]. In addition, LDCT was shown to be cost-effective in health economic evaluations [6,7].

RCTs are the most rigorous type of study but also the most expensive, and they may not always be feasible or appropriate. And because assessing the ultimate value of diagnostic imaging is challenging, radiologists, manufacturers, and other stakeholders need a common understanding of what "value" means and how it can be demonstrated.

In 1991, Fryback and Thornbury [8] developed a hierarchical model of efficacy for appraisal of diagnostic imaging, consisting of six levels (Table 1). Building on this model, Gazelle et al [9] created a new framework 20 years later to provide guidance for assessing the value of diagnostic imaging from a payer perspective. The framework includes the size of the at-risk population,

[a]Siemens Healthcare GmbH, Erlangen, Germany.

[b]Siemens Medical Solutions USA, Inc, Malvern, Pennsylvania.

[c]Siemens Healthcare GmbH, Forchheim, Germany.

Corresponding author and reprints: Sebastian Schmidt, MD, Siemens Healthcare GmbH, Siemensstr. 1, 91301 Forchheim, Germany; e-mail: sebastian.s.schmidt@siemens.com.

The authors have no conflicts of interest related to the material discussed in this article.

1546-1440/16/$36.00 ■ http://dx.doi.org/10.1016/j.jacr.2016.05.013

Table 1. Hierarchical model of efficacy for appraisal of diagnostic imaging

Level	Examples
(1) Technical efficacy	Resolution, sharpness, etc
(2) Diagnostic accuracy efficacy	Sensitivity, specificity, area under ROC curve
(3) Diagnostic thinking efficacy	Impact on diagnosis, change in differential diagnosis
(4) Therapeutic efficacy	Patient management, choice of therapy
(5) Patient outcome efficacy	Morbidity, mortality, QALYs, etc
(6) Societal efficacy	Cost-effectiveness from societal point of view

Note: Adapted from Fryback and Thornbury [8]; the term "efficacy" reflects performance of an imaging device under ideal conditions, whereas the term "effectiveness" reflects performance under real-world conditions. QALYs = quality-adjusted life years; ROC = receiver operating characteristic.

the anticipated clinical benefits, and the potential economic impact. For example, it is suggested that diagnostic imaging technologies affecting large numbers of patients that have a small expected clinical benefit, or hold the potential to substantially increase costs, should require more extensive outcomes data.

The framework by Gazelle et al can be used as a guide for the extent of research needed for a particular imaging technology. We believe, however, that radiologists, manufacturers, and other stakeholders need guidance from which specific strategies can be developed, for example, to establish a diagnostic procedure for clinical use. Our aim is therefore to propose a matrix that allows classification of imaging devices into distinct categories. We use examples to illustrate the application of our matrix.

EVIDENCE AND VALUE

Readers should note that the terms "technology" and "procedure" are used interchangeably at times in this article. As regards the introduction of innovative technologies, it is important to recognize that there is a distinction between evidence and value: The value of a product and the quality of the evidence demonstrating the product's value are independent factors that together affect adoption [3]. Evidence comes from the Latin word *evidentia*, meaning "clearness," and value comes from the Latin word *valere*, meaning "to be worth." Different study designs provide different levels of evidence (or clearness). According to the Oxford Centre for Evidence-based Medicine [10], study designs can be grouped into six levels of evidence, where 1 is the highest level and 6 the lowest: (1) meta-analyses of RCTs, (2) individual RCTs, (3) prospective cohort studies, (4) retrospective case-control studies, (5) case series, and (6) expert opinions. In general, the higher the evidence level, the greater the extent to which a study minimizes systematic error (or bias).

According to Fryback and Thornbury [8], the "imaging process" (which involves an imaging device such as a CT that records an image, usually interpreted by a radiologist) is embedded in a larger "clinical process" (whereby a clinician makes a diagnosis and treatment decision). That clinical process, in turn, is part of the wider health care system whose goal is to improve patient outcomes and reduce costs. The impact of the imaging process on the clinical process and the health care system can be described as value. In other words, the value of an imaging device is defined by what it is worth in terms of improving diagnosis, treatment decisions, patient outcomes, and health care costs. Moreover, considering that imaging devices usually have multiple applications, it can be argued that their ultimate value is some weighted average of their applications [11].

When national coverage decisions are made, the value of imaging technologies is assessed at the highest level of evidence available in determining whether a diagnostic procedure is reasonable and necessary. Noncoverage decisions can be justified by a lack of evidence and/or value.

EVIDENCE VALUE MATRIX

As shown in Figure 1, our proposed two-dimensional matrix combines the evidence hierarchy on the y axis with the efficacy model by Fryback and Thornbury [8] on the x axis. Just like other well-known matrices (eg, product portfolio matrix [12]), the matrix allows classification of imaging devices into distinct categories from which specific strategies can be developed (Figure 2):

- "Stars": High value demonstrated in large prospective cohort studies or, preferably, RCTs. The goal is to establish diagnostic procedures for clinical use and obtain reimbursement (value and evidence should in principle be sufficient to justify coverage decisions).
- "Candidates": High value demonstrated in retrospective case-control studies and smaller case series. These diagnostic procedures potentially have a positive effect on patient management and outcomes, making them strong candidates for prospective studies.