Developing Evidence of Effectiveness and Value: Lessons from Cardiac Computed Tomography Angiography (CCTA)

Over the last decade, researchers, manufacturers and clinicians have invested significant effort in perfecting the use of cardiac computed tomography angiography (CCTA) as a non-invasive alternative to angiography, which is currently accepted as the “gold standard” technology to diagnose coronary artery disease (CAD). Although use of the technology has become widespread, the evidence base to guide that use remains incomplete and clinical recommendations contradict one another. This lack of evidence reflects both methodological complexities in studying the effectiveness of diagnostic technologies like CCTA as well as the early decision by many insurers to cover the procedure. Reimbursement for CCTA was so complete by 2008 that the Centers for Medicare and Medicaid Services (CMS) abandoned a proposed national policy that would have covered procedure only as part of clinical research intended to fill the substantial gaps in evidence.

In the fall of 2008, the Center for Medical Technology Policy (CMTP) convened a “think tank” of 43 experts and stakeholders to discuss lessons to be learned from the CCTA experience. The backdrop for this meeting is the imminent emergence of yet newer cardiac imaging technologies and a major new national initiative to compare the effectiveness of alternative technologies that intended for the same purpose. To what extent, what insights does CCTA offer about building a sufficient body of evidence to guide the use of future technologies?

Background

Coronary artery disease (CAD) or coronary heart disease is a condition in which cholesterol, calcium and other substances (collectively referred to as plaque) build up inside the arteries that supply the heart with oxygen-enriched blood. The build-up of plaque is also referred to as atherosclerosis. Plaque build-up restricts the flow of oxygen-rich blood to the heart, which can cause a feeling of pressure or pain in the upper body, referred to as angina. A completely blocked artery is a heart attack, which can damage the heart muscle or cause death if not treated quickly. Individuals with long-term CAD can develop heart failure, in which the heart is unable to pump sufficient blood to other parts of the body, or arrhythmias, in which the heart beats irregularly. In 2006, CAD caused an estimated 425,000 deaths, with a total of 6.4 percent of individuals over 18 years old (about 14 million people) reporting that they have CAD. Chest pain, the primary symptom of CAD, brings about 6 million people to hospital emergency rooms each year. With more than 60 percent of hospitalizations for chest pains, costing over $8 billion each year, estimated to be unnecessary, better and cheaper initial diagnostic tools could have significant implications for both health care costs and quality.

The standard method of diagnosing CAD is invasive coronary angiography in which a catheter is threaded from the groin or arm to the coronary arteries to inject a radio-contrast dye that illuminates blood flow using x-ray technologies. Because angiography carries risks to the patient, clinicians also use noninvasive diagnostic techniques including electrocardiograms.
EKGs), echocardiograms (ECHOs), and single photon emission computed tomography (SPECT). However, none of these tests provide a direct image of blocked coronary arteries or a definitive diagnosis. Because of recent advances in computed tomography technology that allow for faster images with much greater resolution than was possible in the past, clinicians and researchers have become interested in using CCTA as a tool to diagnose CAD. CCTA is an outpatient technique in which patient is administered a contrast dye through an IV usually placed in a vein on the arm. The CT machine then takes a large number of x-ray “slices” of the coronary system from various angles and compiles them into an overall scan using internal software. While earlier CT machines provided only 16 or 32 slices per image, the current machines provide 64 slices with some scanners providing as much as 256 or 320 slices. The 64 slice machines allow the entire scan to be completed in about 8 seconds thus minimizing distortions due to movement. They also require less contrast dye, which reduces the risk of the patient experiencing an adverse reaction.

Although there are multiple potential clinical scenarios in which researchers could hypothesize a comparative benefit of using CCTA, most discussion has focused on its potential role in diagnosing CAD in two cases: (1) patients in an emergency department presenting with acute chest pain but negative results using one or more of the other non-invasive tests such as EKG or ECHO; and (2) patients with stable chest pain in an outpatient setting.

Proposed Coverage With Evidence Development (CED) for Medicare

As 64-slice scanners gained widespread adoption after 2005, so too did the use of CCTA. A key factor in this diffusion may have been the proliferation of published studies examining CCTA’s potential role in diagnosing CAD. Not all of this research has been based on the newer 64-slice scanners. Studies using the older 16-slice scanners have generally shown no comparative benefit of using CCTA. Among those that did use the newer machines, some have concluded that CCTA can diagnose CAD with sufficient accuracy to warrant its use. However, most were conducted using convenience samples and involved a single clinical site. The lack of evidence from randomized controlled trials or other free of bias and with sufficient statistical power to find a potential benefit left some observers skeptical about the use of CCTA.

As cardiologists and other physicians began to use CCTA, insurers had to face the question of whether the service was medically necessary and thus covered. A major component of medical necessity for CCTA was whether it was a proven (versus experimental) technology in diagnosing CAD. Given the prevalence of CAD in the population, this became a significant question for the Nation’s largest insurer, Medicare. In most cases, Medicare coverage decisions are made by the private contractors, usually insurance companies, who process and pay claims on behalf of the program. Each locale has one contractor that processes in-patient hospital claims (Part A) and another that processes claims for outpatient services (Part B). Hence, the possibility that coverage policy for a given service could vary across the country is built into Medicare program. In some cases, the Centers for Medicare and Medicaid Services (CMS) may choose to review the scientific evidence for the effectiveness of a given service and issue a National Coverage Decision (NCD) that all local contractors must follow.

In considering whether to consider whether to make an NCD, CMS often supplements its own
expertise with a technology assessment of the service’s effectiveness. In 2006 CMS commissioned the Duke University Center for Clinical Health Policy Research Evidence-based Practice Center (EPC) to conduct such an assessment of non-invasive imaging technologies for CAD. The resulting report included a review of the evidence for CCTA.\textsuperscript{10} In June 2007, CMS initiated a national coverage analysis of CCTA. On the basis of the Duke assessment, input from an outside advisory committee, its own internal research, and public comment, CMS issued a proposal on December 13, 2007 for an NCD to cover CCTA for patients with either of two conditions enrolled in approved studies designed to produce credible, definitive evidence about the technology’s effectiveness. CMS first introduced this concept, called coverage with evidence development (CED), in July 2006.\textsuperscript{11} The two conditions for which Medicare coverage of CCTA would be allowed were:

- Symptomatic patients with chronic stable angina at intermediate risk of CAD;
- Symptomatic patients with unstable angina at a low risk of short-term death and intermediate risk of CAD.\textsuperscript{12}

All other uses of CCTA to diagnose CAD would not have been covered by Medicare. In order to be approved under the CED, the studies would be required to address one of three questions:

- Does CCTA have the ability to diagnose or exclude coronary artery disease as well as invasive angiography?
- Does CCTA reduce the need for invasive coronary angiography?
- Does CCTA improve health outcomes for patients with acute chest pain who present in the emergency room or other setting?

The studies would have also had to meet a number of specified standards intended to assure the appropriateness of the protocol as well as the scientific credibility and clinical usefulness of the research results.\textsuperscript{13}

During the 30-day public comment period following publication of the proposed the NCD allowing only CED, CMS received several hundred responses from clinicians, manufacturers, patients, researchers, and payers. All but a very few advocated against the proposal and in favor covering CCTA as a diagnostic tool for CAD without restriction. They countered the stated scientific rationale for the proposed CED with alternative conclusions drawn from the same evidence base used by CMS, additional studies, or personal experience.\textsuperscript{14}

On March 12, 2008, CMS published its final decision, which reversed its earlier proposal for an NCD allowing only CED. CMS decided instead that “no national coverage determination on the use of cardiac computed tomography angiography for coronary artery disease is appropriate at this time and that coverage should be determined by local contractors through the local coverage determination process or case-by-case adjudication.”\textsuperscript{15}

Prime among the reasons why CED was ultimately not feasible for CCTA was its rapid diffusion after 2005. Although this trend was accompanied by a rise in research studies about CCTA, none were large (multi-center), randomized controlled trials (RCTs), which are considered the “gold standard” for clinical research. Because some private health plans and most regional
Medicare contractors covered the CCTA at the time CMS issued its proposed NCD, restricting the service to patients enrolled in approved research studies was seen by patients, providers and manufacturers as removing a benefit that most Medicare patients already had. Widespread coverage of CCTA by insurers also supported the view of some experts and stakeholders that the technology had become standard practice in diagnosing CAD among some patients, likely making the recruitment of patients for an RCT in which they may not receive CCTA more difficult. Further confusing the debate, some of those opposing the proposed NCD cited new research that, because of the time lag between the completion of research studies and publication of their results in peer-reviewed journals, had not been subjected to the scrutiny of systematic reviews like the one Duke University had prepared for CMS.16

Unresolved Evidentiary Issues

In addition to leaving uncertainty about the value of this cardiac imaging procedure, the history of CCTA illustrates several unresolved issues concerning the development of evidence to guide the use of future cardiac imaging technologies and, in some cases, new diagnostic procedures more generally.

- **Moving beyond measures of diagnostic accuracy to measure health and quality of life outcomes.** Researchers traditionally define the effectiveness of a diagnostic test in terms of its performance.17 Although the ability of a test to safely differentiate between healthy and sick patients is an important component of its effectiveness, a true comparison would measure whether the test leads to better health outcomes for the patients (i.e., taking treatment and management of the diagnosed condition into account over time). A new test that offers better performance than the standard diagnostic procedure, but less or no improvement in patient outcomes would not be worthwhile, especially if it costs more than the standard test. However, gathering evidence about patient outcomes associated with alternative diagnostic procedures is more complicated, time-consuming and expensive than is measuring a test’s diagnostic performance. Such research takes more time, sometimes substantially more time, and there can be more variables for which to control, depending on the study design. For new cardiac imaging techniques researchers and policymakers will need to decide when data on test performance is enough, when it is necessary to directly measure health outcomes over time, and if there are intermediate, proxy measures that are reliable predictors of ultimate patient outcomes.

- **Specification of the technology.** When measuring the effectiveness of a cardiac imaging technology, the exact definition of that technology can change over time, especially when measuring effectiveness in terms of downstream health outcomes. For any given diagnosed condition, treatments can vary across providers and evolve over time. When measuring health outcomes, it may not be possible to disentangle the effectiveness of the diagnostic technology from the management and treatment of the diagnosed condition. This possibility can be compounded by any lack of effectiveness evidence for alternative treatments, any evolution of the diagnostic technology over time, changes in clinicians’ confidence in the technology, and the tendency for the clinical scenarios in which the diagnostic technology is used to widen over time (also known as “indication creep.”)
Clinical studies of future cardiac imaging technologies will need to account for these potential difficulties in specifying the exact diagnostic technology being evaluated.

- **Measuring comparative value in addition to comparative effectiveness.** In addition to evidence about comparative effectiveness, policymakers may want to base clinical and coverage decisions on comparative value. If two tests are equally effective, but one costs more, the other test may be judged more valuable and hence, preferred. One component of value is cost effectiveness – i.e. the cost per extra year (or quality-adjusted year) of life added on average. Other factors such as convenience for patients and providers, can also affect one’s judgment of value. The inclusion of cost and value in the analysis of medical technologies such as cardiac imaging has been a controversial issue in the United States. In addition, methods to measure value also vary. There currently is limited guidance and experience to resolve these issues in the development of evidence for cardiac imaging.

- **Timing and role of RCTs.** As mentioned earlier, RCTs are considered the “gold standard” method for developing credible evidence of effectiveness. However, they can prove to be impractical or infeasible research methods, especially for quickly diffusing medical device-based technology. Because RCTs rely on randomizing patients between the standard and new procedure, researchers have an ethical obligation to assure the safety of the new procedure before starting the trial. In addition, generalizability of results usually requires enrollment of patients from multiple, representative institutions. In the case of CCTA, the diffusion of the necessary medical devices and decisions to pay for the procedure occurred before multi-site RCTs could be designed, funded or completed. For future cardiac imaging technologies, the optimal time to begin RCTs remains unclear. Guidance to resolve this question in the future would be useful to researchers and policymakers alike.

- **Role of alternative study designs.** Because RCTs are seen as the “gold standard” for the development of credible evidence, less attention has been paid to the potential contributions of alternative, potentially more practical sources of evidence such as cohort studies, registries, and other types of observational research. Experts have discussed the potential usefulness of practical trials and related study designs more fully elsewhere. To the extent that these methods could produce evidence about the effectiveness of a promising, emerging cardiac imaging technology before it is widely used and reimbursed, the research community should consider their potential early in the technology’s development.

- **Role and feasibility of CED.** Similarly, the case of CCTA suggests the circumstances in which CED is appropriate and feasible for cardiac imaging technologies, if not more generally, needs greater attention. In the case of CCTA, the national CED proposal by CMS came too late.

- **Resources and infrastructure.** The development of clinical evidence for new cardiac imaging and other diagnostic technologies require a scientific infrastructure and resources. CED decisions are potential source of research funding, to the extent they are
appropriate and viable. However, given that most scientific research takes place under the auspices of academic medical centers where RCTs are the preferred means of producing credible evidence, there are fewer funds and fewer researchers for practical trials and research studies outside of academic settings. To the extent that effectiveness evidence are sought in community settings and through methods other than RCTs, policy makers will need to consider how to assure appropriate resources for this work.

Recent Efforts to Improve the Evidence Base

In the aftermath of CMS’s attempt to allow only CED, researchers have designed studies intended to answer questions about the value of CCTA through RCTs. Two multi-center RCTs currently underway, The Coronary Computed Tomography for Systematic Triage of Acute Chest Pain Patients to Treatment (CT-STAT), and the Rule Out Myocardial Ischemia/Infarction Using Computer Assisted Tomography (ROMICAT II) trials have enrolled 750 low-risk and 1,000 intermediate risk patients in emergency department settings to compare clinical and economic outcomes for CCTA to myocardial perfusion SPECT imaging (MPS) as the current standard of care. Two questions about these studies are whether they will possess the statistical power to differentiate the value of CCTA for different types of patients and whether the results of the studies will be generalizable to the population at-large.

By contrast, cardiologists at the Cornell University Medical Center have proposed a very large multi-center, randomized comparative effectiveness study with a parallel registry dubbed the Functional or Anatomic or Both Functional and Anatomic Testing in Symptomatic Individuals Undergoing Evaluation my myocardial perfusion SPECT imaging (MPS) or CCTA, Costs and Clinical Outcomes (FABULOUS) trial. The overall goal of this study would be to compare the clinical effectiveness, cost effectiveness, quality of life and safety in symptomatic patients without known CAD undergoing initial diagnostic coronary artery evaluation by CCTA or MPS. It would enroll 15,000 patients and measure both clinical and cost outcomes over a 12-month period. The National Heart, Lunch and Blood Institute (NHLBI) has considered, but has not, to-date, funded the FABULOUS study.20

NHLBI also held a workshop July 21-22, 2008 to develop research strategies and specific actions to improve the evidence base for evaluating the preventive and therapeutic value of cardiac imaging technologies. The workshop designed clinical trials to assess the value of imaging strategies in four different clinical scenarios and to make recommendations for future research priorities. A peer-reviewed article reporting workshop results is currently in press at the Journal of the American College of Cardiology.

Although the results of these current and proposed studies could resolve many of the questions about the value of CCTA, they come too late. There currently is little sound evidence about the potential value of CCTA in improving patient outcomes that patients, providers and payers can use to guide decisions. Nonetheless, clinicians are using CCTA and insurers are paying for it. In addition to making the recruitment of patients to RCTs difficult, insurers would face opposition if they were to decide not to pay for the procedure in the future because evidence establishes scenarios in which CCTA offers no net value or creates harm. Furthermore, once
CCTA is paid for in one or more situations, it may be more likely to be paid for in all situations, even if there is no evidence to support coverage for all indications.

Concluding Thoughts

The history of CCTA has had important implications for both its use by clinicians. CCTA currently guides decisions about the management and treatment of CAD for millions of patients. However, the value of this form of imaging remains a subjective judgment because the focus of research has largely been on the performance and accuracy of the test, rather than health outcomes. Effectiveness research has also faced multiple methodological challenges. Current guidelines for CCTA conflict with one another, which may lead clinicians to use the technology inconsistently. This case also has implications for other, emerging cardiac imaging technologies such as cardiac molecular imaging technology (CMIT) or cardiac nuclear SPECT. It points out the need to develop the evidentiary framework for studying these diagnostic procedures so that evidence of their comparative effectiveness and value can be available to policy makers and clinicians in a timely manner. Furthermore, the case of CCTA has important implications for the $1.1 billion in new comparative effectiveness research funds appropriated by the American Recovery and Reinvestment Act (ARRA) of 2009 (P.L. 111-5). In order to maximize potential returns for this investment, researchers and policymakers will need a better consensus of what evidentiary standards are sufficient to recommend the use of a given technology as well as a deeper understanding of how to create that evidence. For cardiac imaging technology, development of such evidentiary guidance would be a logical next step.

9 Among the circumstances that can lead to an NCD are (1) conflicting contractor policies for a service; (2) the service is a significant medical advance and no similar service is covered by Medicare; (3) there is substantial scientific controversy about the effectiveness of the service; (4) the service is covered but widely considered ineffective or obsolete; (5) there is evidence of significant under- or overutilization of the service; (6) CMS receives a request for the NCD from an outside individual or organization. Federal Register. April 27, 1999. Vol 64. No. 90 pg. 22621.
A test’s performance is usually measured in terms of its specificity (probability of a negative test result for someone who does not have the condition), sensitivity (probability of a positive test for someone who does have the condition), positive predictive value (probability of someone having the condition if they receive a positive test result), false negative rate (probability of a negative test result for someone who does have the condition), false positive rate (probability of a positive test result for someone who does not have the condition), and the risks of harms from the test itself.


