Trends-in-Medicine

March 2009
by D. Woods

Quick Pulse

Trends-in-Medicine
Stephen Snyder, Publisher
2731 N.E. Pinecrest Lakes Blvd.
Jensen Beach, FL 34957
772-334-7409 Fax 772-334-0856
www.trends-in-medicine.com
TrendsInMedicine@aol.com

MEDCAC MEETING ON GENETIC TESTING

Baltimore, MD
February 25, 2009

A Medicare advisory panel told Medicare that genetic tests are more complicated than other diagnostic tests and need tougher proof in order to qualify for coverage. The Medicare Evidence Development and Coverage Advisory Committee (MedCAC) told the Centers for Medicaid and Medicare Services (CMS) that the evidence needed for diagnostic ***genetic*** tests is different from that of diagnostic testing in general, and that genetic testing resulting in direct patient-centered healthcare outcomes such as mortality, functional status, and adverse events should be required for a CMS coverage decision.

The panel also agreed that:

- Ethical issues relating to genetic tests should not affect the methodological rigor necessary for genetic test trials. Ethics are very important but should not be used as an excuse for inferior trials.
- The ACCE (Analytic validity, clinical validity, clinical utility, and the ethico-legal issues) genetic testing framework and EGAPP (Evaluation of Genomic Applications in Practice and Prevention) guidelines should be used as a basis for future trials.
- Rigorous validation of genetic diagnostic tests is necessary.

CMS asked MedCAC to hold a public meeting on what kinds of evidence CMS could use to decide whether genetic testing, as a laboratory diagnostic service, improves health outcomes. A second meeting, tentatively scheduled for May 2009, will focus on using genetic testing for screening. The Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS) of the Department of Health and Human Services defines genetic testing as "any test performed using molecular biology methods to test DNA or RNA, including germline, heritable, and acquired somatic variations." Medicare may cover a diagnostic test used by the patient's treating physician to guide the physician's diagnosis and treatment. This is different from a screening test, which is used to identify an occult condition or state in an asymptomatic person.

The 15-member panel did not come to many clear conclusions, saying that the questions posed by CMS overlapped, with numerous different possible answers. Panel chair Dr. Barbara McNeil, a radiologist and professor at Harvard Medical School's Department of Health Care Policy, summarized the answers to each question.

QUESTION 1. Are the desirable characteristics of evidence for diagnostic genetic testing different from the desirable characteristics of diagnostic testing in general?

Panel members agreed that both the ACCE genetic testing framework and the EGAPP guidelines are valuable tools and could be used as a basis for future trials. Dr. McNeil said, "What I'm hearing is that, in general, a framework for testing is a framework for testing, with room for the…ethics questions."

QUESTION 2. What are the desirable characteristics of evidence for determining the analytical validity of genetic diagnostic tests?

The panel seemed to agree that the most rigorous tests would be those that meet the gold standards of approval in clinical settings, validated by a group like the American College of Pathologists or approved by a regulatory body such as the FDA. The panel generally liked the idea of adopting some of the EGAPP criteria regarding determining analytical validity of the tests.

QUESTION 3. Beyond aspects of analytical validity, are there meaningful differences in the desirable/necessary characteristics of evidence about the effect of genetic testing on outcomes?

Again, the panel turned to the EGAPP framework on the question of diagnostic assessment. It also suggested adding follow-up in the prognostic assessment question. The panel discussed the differences between using genetic tests for incremental value compared to substantive value. Dr. McNeil said, "This is probably one of the most complicated areas, looking at the incremental value of a test in terms of prognostic ability. And that should definitely be there. Recognizing that is an important consideration…We're talking about patients who change…as a result of reclassification…The other point is that, for common diseases, when we're looking for prognostic factors, we're going to want more studies than for rare diseases."

Regarding pharmacogenomic assessment, Dr. McNeil said, "We're still trying to decide what kinds of clinical studies that you would do – to look at a genetic test in terms of pharmacogenomics…Obviously, we're closely linking the test, choice of drug, and outcome of drug. Is the type of trials we do the same as for anything else?"

QUESTION 4. How confident are you that methodologically rigorous evidence on the outcome is sufficient to infer whether or not diagnostic genetic testing improves patient-centered health outcomes?

The only votes of the day were for this question, and the panel voted only three times, instead of the requested nine – one for each type of assessment. After a circuitous discussion, Dr. Steve Phurrough, director of the coverage and analysis group for CMS, finally boiled down the question, "If you make the assumption that there is good evidence, however you define it, that a genetic test – a diagnostic genetic test – changes physician-directed patient management, is that sufficient for coverage? Most of the evidence is bad. Even if it's methodologically good evidence, but the outcome is only physician-directed patient management changes, should you use only physician-directed patient management in making your coverage decision?"

The panel generally agreed that changes in physician-directed patient management alone are not sufficient for coverage. The vote was mixed (3s and 4s) when it came to indirect or intermediate healthcare outcomes, but generally agreed that direct patient-centered health outcomes are sufficient for coverage.

QUESTION 5. Are there ethical issues particular to genetic testing that may alter the methodological rigor of studies of genetic testing?

The panel agreed that there are great concerns about ethical issues related to genetic testing, including privacy, patients who might be in danger of losing their insurance if certain genetic information were to be made public, and how to deal with relatives of people who are found with certain genetic markers. However, the panel also agreed that companies asking for approval to conduct trials should be held to the highest standards and should not be allowed to use ethical questions as an excuse to be treated leniently.

QUESTION 6. Does the age of the Medicare beneficiary population present particular challenges that may compromise the generation/interpretation of evidence regarding genetic testing?

Although the panel generally agreed that most Medicare beneficiaries would not have much use for many genetic tests (e.g., hereditary disorders), some genetic tests could prove very useful, for example, tests that may show intolerance to some drugs.

BACKGROUND

CMS epidemiologist Dr. Jeffrey Roche presented the background, "There are many tests available for diagnostic use, and each answers a question. But the way that physicians decide what the meaning of the lab test is, and, in many ways, the value, is its effect on being able to provide good information. We recognize that genetic testing raises new challenges …We know that these kinds of tests are here, they're here to stay, and many laboratories…indicate that these tests are widespread, are being used in clinical medicine, and our challenge is to look at how they will help in patient care… Professional societies are increasingly looking at evidence also to find out if genetic testing might be valuable for physicians in practice, the American Society of Clinical Oncology, or the American Society of Chest Physicians, for example. The value to such general approaches…can lead to a consistent and clear approach…to genetic testing."

Dr. Thomas Trikalinos from Tufts-New England Medical Center presented a technology assessment of selected pharmacogenetic tests for non-cancer and cancer conditions. Sixty-two tests were identified in the 2006 Evidence-Based Practice Center (EPC) Technology Assessment reports on genetic tests for patients with cancers. Ninety-one tests with high likelihood of applicability to the Medicare population were identified in the 2007 EPC Technology Assessment report on genetic tests for patients with non-cancer conditions. He said that the two studies were the starting point, and that tests were included which had a high likelihood for utilization in the Medicare population. Tests that had a very large number of publications that could not be reviewed within the work assignment period and tests that have recently been or are currently being reviewed at CMS or the Agency for Healthcare Research and Quality (AHRQ) were excluded.

Four pharmacogenetic tests were reviewed: CYP2C9 and warfarin therapy, VKORC1 and warfarin therapy, APOE and statin treatment, and MTHFR and chemotherapy of the folate metabolic pathway.

Dr. Trikalinos described the ACCE evaluation process for genetic tests, and said that ACCE mapped the four elements in terms of the four pharmacogenetic tests. He concluded, "The value of every test is judged by its ability to affect patient-relevant outcomes. Evidence from different studies has to be brought together to answer the overarching questions…We need a framework for diagnostic tests to contextualize the evidence.

One Proposed Framework for Evaluating Genetic Tests

Level	Explanation
1 – Technical feasibility	Does the test perform reliably and deliver accurate information?
2 – Test accuracy	Does the test contribute to making an accurate diagnosis?
3 – Diagnostic impact	Does the test result influence the pattern of subsequent diagnostic testing?
4 – Therapeutic impact	Does the test result influence the selection and delivery of therapy?
5 – Clinical outcomes	Does performance of the test contribute to improved health of the patient?
6 – Societal impact	Cost effectiveness, ethical issues

Dr. Trikalinos said that the ACCE evaluation process for genetic tests is a "framework that breaks into four questions: analytic validity, clinical validity, clinical utility, and ethico-legal issues. He presented some insights from his report regarding CYP2C9, VKORC1, and warfarin:

- Common CYP2C9 variants affect the metabolism of warfarin.
- Rare mutations in VKORC1 are found in familial cases of resistance to warfarin.

Dr. Phurrough concluded, "We have strong evidence of associations with average maintenance dose, unclear associations with major bleeding, or thromboembolism, but no study measured the effects of testing on patient-relevant outcomes... This is an example…It is a particular trial of sufficient quality to draw conclusions. s a single trial quality sufficient to answer a question? For example, with warfarin, we have a number of trials, and doing the sensitivity testing keeps the International Normalized Ratio (INR) within a proper range… Is there sufficient evidence to determine if the patients are better off? Is there less bleeding? Can you put those two studies together and make a conclusion, or do you need that specific study that says, 'We're going to test and find out if your warfarin sensitivity is genetic, and then we're going to find out what your bleeding and thrombosis rate is, not what your INR is.' That brings us to what kinds of evidence we're going to be looking for."

Ralph Coates, PhD, associate director for science, Office of Public Health Genomics, National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP), Centers for Disease Control and Prevention (CDC), presented a report about EGAPP which he called "a unique method to assess the evidence for determining requirements for specific genomic tests. The project focuses on the role of genetic testing for common disorders, including common chronic diseases." EGAPP's steering committee includes members from CMS, FDA, HHS, NIH, and the CDC. The EGAPP evaluation method includes:

- Careful, explicit, specific definitions of disorder, test, and clinical setting.
- Evaluation of accuracy and reliability in detecting genomic markers of interest (analytic validity).
- Evaluation of accuracy and reliability in predicting disorder or phenotype of interest (clinical validity).
- Evaluation of evidence of improved health outcomes, utility in decision-making (utility).
- Assessment of contextual factors.
- Overall assessment of benefits and harms.

The first EGAPP recommendation was in December 2007 for testing CYP450 polymorphism in adults with non-psychotic depression treated with antidepressants (SSRIs). The working group recommendation showed:

- Insufficient evidence for a recommendation for or against use of CYP450 testing in adults.
- In the absence of supporting evidence and with consideration of contextual issues, EGAPP discourages use of CYP450 testing for patients beginning SSRI treatment until further clinical trials are completed.

The EGAPP working group has made only three other recommendations. In ***Genetics in Medicine***'s January 2009 issue, it said that there was insufficient evidence that tumor gene

expression profiling improves outcomes in patients with breast cancer. The working group also said that there was insufficient evidence that UGT1A1 genotyping can reduce morbidity and mortality in patients with metastatic colorectal cancer treated with irinotecan (Pfizer's Camptosar). Finally, the working group found insufficient evidence that genetic testing strategies in newly diagnosed patients with colorectal cancer reduced morbidity and mortality from Lynch syndrome in relatives, although, as Dr. Coates said, "some groups dispute that."

Coates concluded:

- EGAPP methods may assist in determining requirements for a given test.
- EGAPP methods developed from various advisory group recommendations.
  - Use "evidence-based medicine" approaches.
  - Peer-reviewed, published methods.
  - Generally positive response from stakeholders.
  - Peer-reviewed, published reviews of diagnostic and pharmacogenomic tests available as examples.

Dr. Neil Holtzman, an iconoclastic guest panelist from Johns Hopkins Bloomberg School of Public Health, said, "It's gratifying that CDC has taken the ball and made this kind of progress." Talking about the three recommendations made in January by EGAPP, he said, "It seems that it greatly complicates the task that you perform…in the breast cancer segment …What I'm asking…is the problem that you're often dealing with proprietary tests…the trials of which have been supported by the test's manufacturer. And you have a difficult time in terms of limitations of that data in deciding whether they're beneficial. We're going to see more of them. How are we, the public, going to get unbiased information that tells us whether any proprietary tests add something to what we've already done?...The general approach is to recognize that there may not be a single study that provides all the information."

PUBLIC TESTIMONY

Dr. Mitchell Burken, IntegriGuard: Dr. Burken said that the Carrier Medical Directors' New Technology Work Group, "acting within its purely advisory role, wishes to emphasize the critical importance of crafting detailed guidance on personalized medicine (molecular diagnostic testing) evidence construction. We hope that the MedCAC guidance will include at least a clear specification of whether 'reasonable and necessary' can be met with relatively strong clinical validation studies. If so, what are acceptable designs/matrices for such studies and/or outcomes for clinical utility studies, and what are feasible robust design matrices for all stakeholders to embrace?…I would contend that the term predictive assessment is certainly relevant to this group and suggest that the panel clarify predictive test vs. a prognostic test."

Dr. Mary Fowkes of the College of American Pathologists: "Genomic tests are not unlike numerous other laboratory tests …Clinical decisions in use of a test are still guided by well established performance characteristics…Prolonged survival and response to chemotherapy have been found in tumors. As you consider recommendations on evidence requirements, we ask you to keep an open mind. One set of criteria may not be appropriate for all testing situations."

Dr. Bruce Quinn, a senior health policy specialist with the law firm Foley Hoag: "Does a test have incremental accuracy over existing tools or tests? And how much better is it than other tools?"

J. Russell Teagarden, vice president, clinical practices and therapeutics, Medco Health Solutions: "Requirements for genomic testing should permit:

- Data should be drawn from a broad range of healthcare settings and administrative claims databases when analyzed with appropriate methods.
- Flexibility to account for clinical needs that are more or less urgent."

Teagarden told the panel, "Medco research findings use naturalistic datasets, and we recently found that PPIs (proton pump inhibitors) reduce the effectiveness of clopidogrel (Sanofi-Aventis's Plavix). And, with warfarin, we have found that up to one in five people experience clotting or bleeding events within six months of starting warfarin. We have used this kind of data, from natural settings, to test certain hypotheses." He said that randomized clinical trials are necessary but mainly focus on efficacy, "As for warfarin, we still consider the promise of genomic information that suggests that we could do a lot better. And it's getting really cheap; genomic testing looks promising."

David Mongillo, vice president of policy and medical affairs for the American Clinical Laboratory Association: Most of his member laboratories, if not all, perform genetic testing. He cited an article in the ***New England Journal of Medicine*** in February 2009, written by Janet Woodcock, director of the FDA's Center for Drug Evaluation and Research (CDER), that said "in some cases, randomized clinical trials will be needed to determine whether pharmacogenic trials are necessary. In some cases, they won't be needed. Genetic testing is a vital tool in helping determine treatments for individuals, and that can be validated with a range of diagnostic methods. Randomized clinical trials have significant limitations when applied to diagnostic situations… There's unnecessary delay in access to diagnostic tests that could immediately help patients, reduce side effects, and help control healthcare costs. Balance is needed…for solid science and patient access."

Dr. Roger Klein, medical director, molecular oncology, Blood Center of Wisconsin: He said that DNA- and RNA-based tests have a wide range in clinical applications and in

"the furtherance of personalized medicine…(But) some genetic tests lack the gold standard for accuracy and results. Molecular diagnostic methods tend to excel analytically...It is possible that age-related characteristics could potentially interfere with study recruitment…Generalization of results to Medicare patients may not be appropriate."

PANEL QUESTIONS FOR EXPERTS

Deborah Shatin, PhD, of Shatin Associates, a health research consultant, asked about the SSRI study, "The information wasn't sufficient to determine whether a class of drugs was effective for patients in this instance?" Dr. Coates responded, "There were small numbers of patients in the studies. The studies were small to begin with, and when you try to separate them out by medications, there's even less information available. The problem is variability in the metabolism among the SSRIs, so there might need to be tests for each specific SSRI."

Dr. Steve Gutman, professor of pathology, University of Central Florida, said, "I think that EGAPP is the gold standard. It's like driving a Cadillac…but it takes a lot of time to do. You hear about flexibility and contingency. Is it my understanding that your group is looking at ways of dealing with data in a more facile manner, allowing for interim decision making? Dr. Coates answered, "The process has taken a long time – four years and only four recommendations – there is recognition that the evidence reviews are complex and take a lot of time. There's a plan now to get a model for rapid reviews and to essentially address the same six issues, from specifics test, clinical scenario, disorder, through validity, but in a more rapid fashion. It hasn't been clearly worked out yet…We recognize that the process is slow and limited, and I think it is still a difficult issue. All of what's available needs to be taken into account, and it has to be a specific test for a specific use."

Dr. Fowkes and the panel chair discussed Dr. Fowkes' concerns about molecular testing and the 21-gene assay (Genomic Health's Onco*type* DX).

- *Dr. Fowkes:* She said that the company has a website which doesn't indicate breast tumor grades "because of the sparseness and availability of the data. However, current standards for staging…state that the histologic grade (Nottingham) provides a strong predictor of outcome in patients and should be incorporated into breast cancer staging systems, so my concern with some of the genetic testing is the thought that genetic testing is going to give something that is more or better than the pathology evaluation of the specimen, and you have to be sure when looking at the literature."
- *Panel chair:* "We don't want to pick on that one test in particular, but here's the example of what is the incremental value of a test?"
- *Dr. Fowkes:* "But the diagnostic criteria for tumors is changing too…When breast cancer is being evaluated, and tissue is removed and evaluated by a pathologist and a specific tumor stage is made for the tumor, do the tests include ER (estrogen receptor) positivity? In instances where the tumor is ER positive and node negative, the oncologists tend to request that the Onco*type* DX (Genomic Health) test be done. However, the pathologist evaluating the tumor – if the pathologist looks at the grade of the tumor histologically using the Nottingham criteria, the problem of recurrence correlates to that grading and gives similar results and similar findings to the test."
- *Panel chair:* "We're not evaluating Onco*type* DX here. We need to realize that we're talking about issues that relate to the consideration of genomic tests. Medicine does move on, and we have to remember that we won't always have everything lined up like peas in a pod."

Dr. Mina Chung, a cardiologist at the Cleveland Clinic, asked Dr. Coates about the SSRI study, "Where you found on consistent association and clinical validity and that led to insufficient evidence. Given that many of these tests look at single genetic tests and don't account for environmental interactions, is that the type of information that you could include in your contextually considered recommendation? What would you have used to make a negative recommendation as opposed to insufficient evidence?" Dr. Coates answered, "A negative recommendation has to have a balance of negatives and harms and is one of harm overall."

Clifford Goodman, PhD, senior vice president of the Lewin Group, asked Dr. Coates, "If we knew from peer review literature that a group of well-defined patients, through phenotype and genotype – we have evidence through a randomized clinical trial about what treatment works the best, as long as we could get to that point without randomized clinical trials, which it sounds as if we may be able to do so, then the only randomized clinical trial evidence we might need would be for clinical utility. So, I'm asking if we can use non-randomized clinical trials to get that far through the analytical framework and then use good randomized clinical trial evidence for a patient group, is that a useful construct?" Dr. Coates answered, "That could be considered by the working group."

Panel vice chair Dr. Steven Pearson, president of the Institute for Clinical and Economic Review, Massachusetts General Hospital and Harvard Medical School, asked Dr. Coates, "Can you synthesize what are some of the concerns? What are the standards for the kind of evidence that CMS should look for?" Dr. Coates answered, "For many of these lab developed tests, there's only one source of getting it done. So one of the issues CMS might look at is the availability of evidence of the assay of that lab, and that's often not available. Or sometimes not available. Maybe going to the test and requesting that information. That would be reliability. Issues of how well tests perform in different settings and different labs might be less important, but how well they perform in the specific

population for which they are proposed (is of concern)." Dr. Phurrough added, "CMS regulations are reasonable...There is a presumption that the physician would be a reasonable use of the test."

Another panel member commented, "Our job is to help CMS provide a roadmap – ground rules for kinds of evidence that will be expected…for a Medicare coverage decision. Are we headed for trouble in terms of the kinds of evidence that we'd like to get vs. where most of the evidence is for lab-developed tests? Might not they be in the public domain? Lab-developed tests don't go through the pre-market processes other tests go through. Are we going to be asking for evidence that we're going to get? Is that going to be a gap?" Dr. Coates answered, "Yes."

Dr. Marion Danis, chief of the bioethics consultation service at NIH Clinical Center, asked how larger datasets could be generated. Dr. Klein said, "We're dealing with diagnostic tests. It can be challenging to get funding. The NIH likes to fund basic sciences but doesn't like to fund how well tests work. We try to pool resources, but…clinicians don't like to order up tests that don't help them…If it's funded, everybody would love it, but it's hard to put that burden on individual laboratories." Dr. Fowkes added, "There is no uniformity in pathology. Even if you pooled everything, it's small. And to have everyone agree would be very difficult." The panel chair said, "That's probably not an acceptable answer in 2009."

The Medco representative surprised the panel, saying that Medco has a "pretty high percentage" of genetic information on its patients, including patients on tamoxifen and warfarin, "We put it in the patient's profile like allergy information, and if we get information that's relative to the phenotype on another drug, we're able to collect that information…We're getting a sample of whoever is in these programs, and we take it all."

Only at this point in the discussion did a panel member, Dr. Catherine Eng, medical director of On Lok Lifeways, ask about the potential harm from genetic testing. Dr. Coates answered, "I can illustrate the kinds of harms from the reviews and recommendations we've done so far. On recommendations for colorectal patients, the harm identified there was that even though it appeared that specific genomic tests could predict which patients might have more adverse side effects, it was unclear even in that group whether reducing the dose or medication was beneficial to the patient because there was evidence of reduced efficacy of the chemotherapy. The question was of reduced side effects vs. harm (recurrence of cancer) and that wasn't clear from the information then available. On the using of tests for SSRIs, one concern was about a complex situation where the clinician and patient are trying to choose a particular test and then doing the dosing. Adding a test in that situation wasn't a good prediction. One would add cost to the people being treated, and there could be harm because the prediction wasn't good; the prediction could result in giving the wrong dose of the wrong medication."

PANEL DISCUSSION OF CMS QUESTIONS

QUESTION 1. Are the desirable characteristics of evidence for diagnostic genetic testing different from the desirable characteristics of diagnostic testing in general?

Discussion included:

- *Consultant:* "I would pose that we have not heard anything that would push us off the Thornberry framework, and I would love to hear what would push us off that framework."
- *Dr. Gutman, the pathologist:* "I like that framework too, and a test is a test is a test, and where there is any (difference) at all is in the ethical area, where you might damage a patient in terms of their insurability…and a patient's family. A test is a test is a test, and you might be a little more sensitive to that."
- *Public health physician:* "One of the things that surprised me is that there seems to be very little difference here between germline somatic assay and other assays. One is dealing with inherited characteristics and the fact that whether to…make a discovery of a person's genotype has relevance for relatives…That has major concerns that raise issues of conformed consent before one does germline genetic testing, such as in predispositions to cancer… In many genetic tests for common diseases, one is dealing with relatively low predictability. We're not talking about all or none, or yes or no for a test result. We have to recognize that seldom will we have a perfect fit between test results and reality."
- Dr. Phurrough asked Dr. Trikalinos why he thought that the ACCE framework might be valuable. Dr. Trikalinos said, "If you look at the ACCE questions…ACCE is specifically tailored to study genetic testing and has specific questions (in that area)."
- *Dr. Danis:* "It seemed to me that the translation or the ability to translate from the other framework was very helpful and the questions (raised) about germline really reflect many of the ethical issues and don't necessarily undermine the extent to which you could argue that the desirable characteristics of a genetic test."
- *Panel chair:* "What I'm hearing is that, in general, a framework for testing is a framework for testing, with room for...the ethics questions. In general, Thornberry is good…(Dr. Trikalinos) was saying that the ACCE framework provides more specificity…Are we quibbling here? I get the sense that we think they're pretty much the same."
- *Eleanor Perfetto, PhD, the industry representative (Pfizer):* "It appears to apply to any kind of diagnostic testing, so no there isn't any specific characteristic that separates them."

- The consultant said that the Thornberry framework does mention something related to CMS regarding cost effectiveness. Dr. Phurrough said, "Your job is to tell us what the regulations should be and we apply that."
- *Panel chair:* "It strikes me as a little beyond the scope of this panel to talk about cost-effectiveness."
- *Dr. Shatin, health research consultant:* "We're looking at multiple items. Are we saying for a genetic test to meet relevance it has to meet all Level 6 items here? Or do we say we use it as a framework?"
- *Panel chair:* "The latter."

QUESTION 2. What are the desirable characteristics of evidence for determining the analytical validity of genetic diagnostic tests?

Discussion included:

- *Pathologist:* "You want accuracy or trueness; some comparison to some traceable method, working method, and when that's not available you make do with what you can. You always want robustness, precision, proper stresses in the study, and so it's a home roulette test. You want to know the specificity testing – how often interfering materials or substances will cause negative results, and you might want to know the level of quantitation or the level of measurement."
- *CMS's Dr. Phurrough:* "Is the maturity of genetic analysis such that we could be comfortable that labs in general can be assumed to have similar characteristics – the same characteristics as we think should be across the field? Or would some labs be challenged with coming up with accurate results?"
- *Public health physician:* "No. And this raises a very important point. Many of the tests developed today are in single labs…When the tests get out there, and more diagnostic kits that FDA would review…The gold standard of approval of testing in a clinical setting is proficiency testing…One would like any single test to meet standards set by some independent outside proficiency testing program."
- *Panel chair:* "Does that mean that if hospital X comes up with its own test, we should be reluctant to consider the validity of that test in the absence of something by the American College of Pathologists or some other group?"
- *Public health physician:* "Yes."
- *Consultant:* "This in fact is one of the distinctions between ACCE and Thornberry. Thornberry is good at the six levels. ACCE is more detailed, and it reduces to analytical validity, clinical validity, and clinical utility… We may want to state or consider stating or accepting the use of the terminology as important descriptors of evidence as CMS might require. They have been defined for lab tests. Now when you get to each of these three points (i.e., analytical validity)…as our friends at Blue Cross Blue Shield do, one criteria for evidence requirements is that the technology has passed muster with the applicable regulatory authorities, which is typically the FDA, and CMS might want to consider…that tests have passed muster with the applicable regulatory authorities."
- *Dr. Danis:* "I suggest that maybe some of the EGAPP criteria concerning evidence might be something we want to adopt."
- *Panel chair:* "Is there any reason to think that if EGAPP has spent months and years developing new criteria that we could improve them?"
- *Dr. Eng:* "The field is moving very fast, and the question is whether there is flexibility in that process."

QUESTION 3. Beyond aspects of analytical validity considered above, are there meaningful differences in the desirable/necessary characteristics of evidence about the effect of genetic testing on outcomes for three testing paradigms below? If yes, please consider Question 4 separately for each paradigm. If not, please consider Question 4 to apply equally to all three.

a. Diagnostic assessment (e.g., testing for the variant of the gene HD associated with Huntington's disease).

b. Prognostic assessment (e.g., assessment of gene expression in tumor tissue to evaluate likelihood of distant recurrence in patients with early stage breast cancer).

c. Pharmacogenomic assessment (e.g., testing for variants in the K-ras gene which indicated absent response to certain chemotherapy for colorectal cancer, for example, cetuximab).

Discussion included:

- *Dr. Maren Scheuner of the RAND Corporation:* "The distinction between analytical validity, clinical validity, and clinical utility is relatively clear, but is that what I'm hearing?"
- *Dr. Danis:* "It seems to me that we should take the concern about risk associated with a diagnosis and use that as the criterion for how far up or down you go on what level of certainty you demand."
- *Dr. Mark Grant of Blue Cross Blue Shield:* "Test accuracy in and of itself, and how you weigh the consequences of the test, has to do with the benefit of utility… so I'd tend strongly to take that view...the way sensitivities and specificities are set widely vary."
- *Panel chair:* "Could I suggest that we move on to prognosis? We could make some suggestions and here I see something missing – which is follow-up. We talked a lot about short-term and long-term follow-up – that's probably required when we're looking at prognostic tests."

- *Dr. Grant:* "When the test is added for incremental value or substantive value, that question is very important."
- *Panel chair:* "That's probably one of the most complicated areas – looking at the incremental value of a test in terms of prognostic ability. And that should definitely be there. Recognizing that that is an important consideration…We're talking about patients who change …as a result of this classification."
- *Dr. Danis:* "So we'd have to show some evidence of reclassification."
- *Dr. Grant*: "Not just in terms of clinical validity, but also in terms of potential for harm. In fact, they have clinical consequences that are meaningful."
- *Dr. Phurrough:* "The CMS party line here maybe is that we should not attempt to compare what kinds of recommendations you're making here on diagnostic tests into other types of diagnostic testing…We may say we're applying higher standards, but because we don't have standards currently, you should not attempt to determine whether these are too high compared to others."
- *Dr. Grant:* "I'd make a strong case that these standards are not too high. They are minimal, and there are no other means to evaluate benefit and risk without knowing… how treatments are subsequently changed. In EGAPP, it is not informative about classification, but there are other ways to do it. I would argue this is no different than what we'd want to see across the board for any diagnostic test."
- *Public health physician:* "This distinction between diagnosis and prognosis is a little problematic. We're told people already have symptoms of disease. So we're not predicting. We're saying, looking at a person with a constellation of signs and symptoms, a genetic test tells you they have a certain genotype. The value isn't saying whether the patient has the disease, but whether it will help us design the regimen for that person. Will having the genotype influence your decision of what kind of therapy to use?"
- *Dr. Gutman:* "From an FDA perspective, prognosis, predictive testing, and screening tests all are diagnostic. This speaks to the issue of reclassification. While I have no argument against looking at a prognostic test in terms of reclassifying a patient, that should also apply to the original diagnostic test. How that classifies patients might also be interesting."
- *Dr. Eng:* "I have a question about Alzheimer's. I think that we have to think about common conditions. We have to have a higher level of evidence."
- *Dr. Grant:* "Every decision has a different degree of uncertainty accompanying it. It will vary according to multiple factors. One thing missing from here is that a lot of this is all about quantifying uncertainty – providing information to decision makers. The only piece here really has to do with point estimates – but the thresholds of decision making are going to vary according to multiple factors, and I think they should."
- *Panel chair:* "We're on prognosis, and we're talking about clinical validity…When we're thinking about prognosis, we should be emphasizing the incremental value and if that's the case we definitely need reclassification matrices…We need to go marginal increments, and we can only do that if we know how many patients move up and down the spectrum."
- *Dr. Grant:* "When you're introducing a decision whether to use a drug or not, you have to think about the risks inherent in that; it's a more involved question…How accurate is the test to determine if a patient would benefit or be harmed from the therapy? I would say, in general, that the same principles apply."
- *Panel chair:* "Can you trust case by case studies?"
- *Dr. Gutman:* "I think that, when you link a drug to a diagnostic test, that you certainly immensely raise the stakes because the drug becomes a slave to the diagnostic, and if you choose the wrong patient, you have an impact on the efficacy of the drug. You've complicated the design, and in some ways, you've telescoped the design… The fight in the genomics area relates to whether we need to look at entire population, do all comers studies, or use feasibility data and background data, etc., and study only biomarker relevant patients. The disadvantage is that at the end of the study you only know that the drug worked in biomarker relevant patients. You don't know what the drug was like in patients who were excluded."

QUESTION 4. For each type of outcome below, how confident are you that methodologically rigorous evidence on the outcome is sufficient to infer whether or not diagnostic genetic testing improves patient-centered health outcomes? (For each lettered outcome type, assign a number from 1 to 5 – lower numbers indicate lower confidence)

a. Changes in physician-directed patient management.

b. Indirect or intermediate healthcare outcomes (e.g., changes in lab test results such as hemoglobin or time to achieve a target value).

c. Direct patient-centered healthcare outcomes (e.g., mortality, functional status, adverse events).

Discussion included:

- *Dr. Phurrough:* "This is a general discussion about whether you think changes in physician-directed patient management is a good indicator, and whether it differs by diagnostic, prognostic, or pharmacogenetic."
- *Dr. James Puklin, an ophthalmologist from Kresge Eye Institute at Wayne State University School of Medicine:* "That's a problem with American medicine. The majority of people are not being properly managed. How can you

expect that anything is going to be influential in private practice? The fallout is incredible after you get outside of academic areas."

- *Dr. Grant:* "It would be wonderful if you had random clinical trials, but that doesn't happen for the obvious reasons. On the other hand, we're dealing with an indirect body of evidence, but ultimately it can be explicit or implicit. We have accuracy data, intermediate outcomes, ultimate clinical outcomes, and where do we stop along the way? But we're not grappling with analytical validity, clinical validity, and clinical utility, but we need a greater degree of explicitness not only in linking the pieces of puzzle, but also in degrees of uncertainty. The more you can tie it in to a hard outcome, the better off you are. It might be worthwhile to consider taking a more explicit approach."
- *Dr. Holtzman:* "Part of the problem for me is this diagnostic, prognostic, and pharmacogenic – within the category of diagnostic tests – there are some that have prognostic significance. Within prognostic, some will influence therapy and some will not. That sort of overlapping classification is better than simply the three because they are overlapping."
- *Panel vice chair Dr. Pearson:* "We are looking for some tangible evidence that changes the pathway of care, so classification comes up critically important for clinical validity, and I do think that in general, in a diagnostic situation, if we can have a very clear set pathway of care that's established for a subgroup, then I don't necessarily think we have to see further study. On the other hand, warfarin is interesting, and there could be a new classification. Right now we treat everyone in the middle with an algorithm. Some will receive higher doses upfront, and some will receive lower doses. We may need a randomized clinical trial to see how the patient management changes due to the outcome. Reclassification is an absolute floor for validity."
- *Dr. Puklin, the ophthalmologist:* "Regardless of the tests, I think that the changes in physician-directed patient management is going to be directed by where the tests are used and where the patients are treated. Patients sent to cancer care centers, state of the art centers, will get genetic testing, and it will reduce mortality and morbidity. But if it's warfarin testing to bring the INR to a proper level, I can tell you…patients away from academic medical centers are not being monitored correctly presently on coumadin or warfarin…So, the issue about physician-directed patient management depends on where the test is going to be used the most. This is a highly variable situation."
- *Dr. Holtzman:* "It's very difficult to make general statements…This ABC classification is overlapping…It seems that if MedCAC and CMS are really concerned about validating reimbursements of genetic tests, one has to take a different view of classification. Many things are underlying what I'm saying; one cannot look for generalizations. Genetics itself is the study of variation, and one has to set up some sort of hierarchy as to the stringency of what is required."
- *Dr. Chung, the cardiologist:* "What is desirable may not necessarily be what is achievable. I am stuck on clinical validity vs. clinical utility issues. There's only one warfarin study that applies to clinical utility, zero in the others…Although, yes, it would be great to have these studies as evidence of support for changes in physician-directed management."
- *Panel chair:* "I've heard a couple of things. One is that generalization is not possible. Did I also hear that it's very difficult to answer these three questions, but if we had to, the most important endpoint was C, and we would want the most rigorous evidence for C, and we would accept less rigorous for B, realizing there is interaction between B and C. And when we get to A, we get to 'should vs. did.' I feel like we're stuck."
- *CMS's Dr. Phurrough:* "This is fascinating. If you make the assumption that there is good evidence, however you define it, that a genetic test – a diagnostic genetic test – changes physician-directed patient management, is that sufficient for coverage? Most of the evidence is bad. Even if it's methodologically good evidence but the outcome is only physician-directed patient management changes, should you use only physician-directed patient management in making your coverage decision? We have a good example right here. We're about to finish a coverage decision on PET scanning in cancer patients. And the level of evidence has been non existent, except in rare instances. In the last decision we said, collect more evidence. We allowed that to be done in a registry of physicians reporting how the results of the PET scan when added to other diagnostic tests available changed management of the patient. It's all the data we had. We don't have anything that said that patients are better with more chemo, that they are hospitalized less and lived longer, etc. We don't have any of that. But we do know that physicians changed their minds about what they were doing. That's not uncommon; that's the only evidence we have. Physicians are going to do something different because of the test provided. Is that sufficient to decide or should we use only that level of evidence?"

QUESTION 5. Are there ethical issues particular to genetic testing that may alter the methodological rigor of studies of genetic testing?

Discussion included:

- *Dr. Phurrough:* "Are there privacy issues in clinical trials? Would someone do a clinical trial differently because of our concerns?"
- *Consumer representative:* "The answer is yes. What we're seeing in the popular press is a tremendous amount

of distrust. Privacy issues are super important to lay out, and I could imagine a clinical trial being destroyed because of how the data will be used."

- *Dr. Phurrough:* "Does that change the kinds of trials we would accept? If it does, would that alter the answers to the previous question?
- *Consumer rep:* "Yeah, sure."
- *Dr. Danis:* "When you think about the ethics of research …there's the question of generating scientific data that is valuable to society to make therapeutic decisions and whether we're going to spend money for it. I don't think that we should undermine the rigor of science; we might have to do some special things considering confidentiality."
- *Dr. Scheuner of RAND:* "The consent process is more and more rigorous. I haven't seen a case where patients' rights have been trampled…I don't understand the ethical issue maybe."
- *Panel chair:* "Consent aside, the issue is absolute security and privacy."
- *Dr. Puklin:* "There are enormous ethical issues surrounding all of this. Patients are signing all sorts of consents, but in the clinical realm, I've known several people who have a particular gene (BRCA) in their family. In order to find out whether they were positive, and so not to affect employability, they went to Europe in order to have the results hidden from their medical records. The ethical implications and social implications are enormous in that age group (under 65). For the Medicare age group it is not nearly as creepy; these patients have established disease."
- *Dr. Holtzman:* "In evaluating evidence and collecting evidence for studies, they should all be approved by review boards…among clinical labs or for profit labs that may be more of a problem, so that is an issue that should be looked at as one evaluates studies. Now, archived specimens is another issue. You have specimens with identifiers, and their use has not been approved by the person who gave that specimen. The question is whether you can go back to that individual. The third issue is about BRCA 1 testing. It is germline testing, and this doesn't come up so much in the Medicare population – notifying relatives, whether that information about his or her genotype may be given."
- *Panel chair:* "While everybody is consented, except for private labs, we have to worry a little bit extra about privacy. There is the issue of underwriting, and so people could have their insurance status changed as a result. That wouldn't apply to the Medicare population as much as to someone under the age of 65. There is also the responsibility to relatives if an individual is found to have a positive test. What to do about archived specimens? If the patient is dead, it's not an issue. If the patient is not dead, we try to track them down."
- *Dr. Phurrough:* "I'm going to do what I shouldn't do... While there are important issues around what could be done with genetic information, I think that should in no way change how trials are done. Trials should be rigorous. We shouldn't accept a lower status if we are concerned that the data won't be handled appropriately. I also think that it's a bit inappropriate for us to consider that we're going to let some data have less of a control from a privacy point of view and other data needs more control. The biggest issue is the issue of what trials we're actually doing. Are we actually creating the potential for having greater exposure because we don't design trials that don't answer questions that need to be answered? The really ethical issue, perhaps more in genetic issues, is that we ought to be designing the trials up front that answer the questions, and it's not ethical not to design the trials (properly)."
- *Dr. Pearson of Harvard:* "We want to send a message that this should not be an excuse ***not*** to do the trials."

QUESTION 6. Does the age of the Medicare beneficiary population present particular challenges that may compromise the generation/interpretation of evidence regarding genetic testing?

Comments included:

- *CMS's Dr. Phurrough:* "The question is there an age range at which genetic testing may not provide any information because of the age of the patient, recognizing that at some ages some people are not the same age as they are functional. Is there a benefit in a 75-year-old in getting a particular test if their life expectancy is 10 years if the test doesn't offer a therapy that would improve the lifespan, etc.?"
- *Consumer representative:* "I think that health status is much more important than age."
- *Dr. Holtzman, public health specialist:* "I was a little surprised when I was invited to serve because I did not think that there was much genetic testing that would be beneficial to Medicare beneficiaries. If you go back to the genetics of diseases, almost all appear earlier than age 65. Alzheimer's almost always appears before 65."
- *Dr. Phurrough:* "So it shouldn't be part of the 'Welcome to Medicare' plan?"
- *Dr. Shatin:* "I was perturbed by the wording of this question. I'd say that there may be some opportunities for the elderly population, and, for example, you might find out if you are liable to have a severe adverse event, particularly with pharmaceuticals. So that's the flip side."

♦