

Steep climb to suitable reference standards

William Check, PhD

February 2013—It's a long way from ancient Greek philosophers to modern-day clinical laboratory directors. Yet both types of scholars have one thing in common: the pursuit of truth. Socrates and his disciples thought of truth as correspondence to an objective universal ideal in the mind. Today's clinical laboratory scientists need a more concrete standard against which to measure their results, leading to the continuing search for suitable reference materials to be used in method development, test validation, internal QC, assay calibration, and proficiency testing.

For a laboratory test to produce true results reliably, reference materials are required that are well-characterized, homogeneous, stable, traceable, and commutable, Lawrence J. Jennings, MD, PhD, D(ABHI), noted. He was speaking at a plenary session on the lack of laboratory reference materials at the Association for Molecular Pathology 2012 Annual Meeting on Genomic Medicine. The talks of the plenary's three speakers were intended to be complementary, he told CAP TODAY. "I spent most of my time talking about the challenges we face [in PT for molecular oncology testing] and justifying the approach to move away from tissue. David [Barton, PhD] and Lisa [Kalman, PhD] spent most of their time talking about the work they do characterizing [genetic] reference materials and what is available."

Dr. Jennings proceeded from the premise that proficiency testing surveys (also known as external quality assurance, or EQA) should be graded. The CAP's Molecular Oncology Committee, of which Dr. Jennings is chair, has made that a goal. Grading requires determining truth, and that means reference standards. However, "we often don't have suitable reference materials," said Dr. Jennings, director of molecular pathology and of HLA and immunogenetics and assistant professor of pathology at Northwestern University Feinberg School of Medicine and Lurie Children's Hospital of Chicago. As a result, grading may be done by using referee laboratories or by consensus—preferably 80 percent—of reporting labs. "These are not always the best options," he said at the meeting.

Later, in an interview, he elaborated. "Over the last several years we have used tissue [as reference material for PT]. Molecular pathologists like to use tissue because it is compatible with how pathologists work. We can identify the tumor and do macroscopic dissection to enrich for tumor." Now the Molecular Oncology Committee would like to move away from tissue and toward cell lines for solid tumors, using a different cell line for each biomarker analyte—BRAF, EGFR, KRAS, and others. "This trend started several years ago with evaluating hematological malignancies for minimal residual disease," he says. Initially leukemia cells from patient samples were used, but there was difficulty getting sufficient amounts of high-quality material, so they went to cell lines. "We are doing the same with solid tumors now," he says. "But it has been difficult to get pathologists to accept it. People like to see that tissue."

In the AMP session, Dr. Jennings gave examples of the kinds of problems that can arise from using tissue as reference material. In CAP Survey 2010B, testing for KRAS in adenocarcinoma of the colon, two-thirds of the laboratories reported a positive result and one-third reported negative. This discrepancy appeared among committee members as well. Sampling error in sending out the tumor specimens was ruled out. All blocks were from the same patient. However, it was found that the tumor sample was heterogeneous, even within the primary tumor. Data show that eight percent of primary and 31 percent of metastatic colon cancers are heterogeneous for KRAS (Baldus SE, et al. *Clin Cancer Res.* 2010;16:790-799). "So this was likely to happen again," Dr. Jennings told CAP TODAY. "How could we prevent this?" Testing all parts of all blocks for mutations would be "undoable," he says. "This experience helped convince people to move to something more reliable, better-characterized, and homogeneous."

A second example concerned heterogeneity of allelic burden within tissue. A specimen of adenocarcinoma of the lung was sent to 82 laboratories to test for EGFR. Ninety percent reported an Exon 21 mutation; all used real-time

PCR or pyrosequencing. The 10 percent reporting negative results all used Sanger sequencing, which has a lower sensitivity. Was this a problem of limit of detection? To investigate, the committee looked at cellularity. Percentage of neoplastic cells in the lesional area reported by Survey participants showed a broad distribution. "Their numbers were all across the board," he says. The same was true in two KRAS Surveys.

Next they sent photos of H&E-stained tissue from a fairly differentiated adenocarcinoma. They asked, What is the percentage of neoplastic cells in this image? Answers ranged from 30 percent to 90 percent. Similar variability was seen in eight more photo challenges. "Even with a photo people were all over the board," Dr. Jennings says. "We cannot trust a pathologist to report back the cellularity and thus allele burden. So we can't assume all labs are getting the same tumor burden. We can't get past the fact that we don't know what we sent them."

A third problem is poor or inconsistent quality. For instance, a large lymphoma specimen was so degraded that participant laboratories said they couldn't analyze it.

Poorly characterized material can be a problem, such as a sample lacking a lesion. One tissue was positive by FISH break-apart fusion probe but negative by rt-PCR. In this sample, gene rearrangement did not yield expression of the fusion transcript.

Other obstacles are lack of samples with rare mutations, lack of sufficient material, and use of samples with single analytes (single mutations). "Labs are doing more multianalyte testing these days," Dr. Jennings notes, adding, "How can we evaluate all those analytes in one sample?"

The lack of well-characterized reference materials is detrimental not only to PT programs but also to validation during assay development. "We are all struggling to find homogeneous, well-characterized material for validation," he says, speaking as a director of a small molecular lab. "For large molecular laboratories such as the Cleveland Clinic or Mayo, that is not a problem. For us to get sufficient numbers of well-characterized specimens, we rely a lot on directors of larger labs. That is not the best way to do it."

To pursue the goal of cell lines as reference materials for molecular oncology, Dr. Jennings has worked with Lisa Kalman, PhD, of the Centers for Disease Control and Prevention. "For the past eight years or so, Dr. Kalman's group has focused on reference materials for heritable conditions," he says. "They have been very successful in generating panels of materials for such things as Duchenne muscular dystrophy and fragile X." Dr. Jennings attempted in 2011 to get the group to work on molecular oncology. While the Coriell Institute for Medical Research has cell lines for many heritable conditions, there is no repository of cell lines for molecular oncology analytes. Dr. Jennings and Dr. Kalman are working with the National Cancer Institute, American Type Culture Collection, and Coriell to see if they can establish such a repository. "I would like things to move much quicker than they have been," he says.

"As we move away from tissue specimens," he continues, "I am very cognizant that cell lines do not address all aspects of pathology practice, such as microdissection or extraction and amplification of DNA from FFPE [formalin-fixed paraffin-embedded] tissue." One way around this is to include method-based challenges that are not tumor specific or mutation specific, such as a cellularity challenge and an FFPE challenge.

Dr. Jennings has an additional concern. "As we move to multianalyte testing, we may obligate people to do multianalyte challenges. They might not want or desire that. Should we have single or multiple Surveys? One argument is that we should allow people only to pay for what they want." Participants in molecular oncology proficiency testing are now being surveyed to find out how many laboratories are testing for more than one analyte. "We know that a large majority are already doing at least two of three biomarkers," Dr. Jennings says. "There would be a tremendous advantage to combining several analytes into a single Survey."

Sharing the European perspective on reference materials for genetic testing was David E. Barton, PhD, chief scientist and associate professor at the National Centre for Medical Genetics, Our Lady's Children's Hospital, Dublin. He spoke about the work of EuroGenTest (EGT), which is a consortium of 35 partners whose goal is harmonization, validation, and standardization in genetic testing across Europe through training, EQA, and control

materials. EGT sent a survey in June 2010 to 910 labs in 32 countries. Responses came from 291 (32 percent) labs. One hundred ninety-eight labs reported using genetic reference materials—samples of defined genotypes obtained from external sources. Of these 198, the vast majority, 69 percent, obtained reference materials from colleagues, while 46 percent used EQA/PT materials, 40 percent used certified reference materials, and 30 percent used cell lines.

“Even in what we consider a highly developed genetic testing environment, which we have been doing for 20 years or more, still there is this almost casual exchange of materials between colleagues that seems more suited to a research environment than a regulated clinical testing environment,” Dr. Barton told CAP TODAY. “Partly this is because there just aren’t formally designated reference materials available for much of genetic testing.”

Reference materials are most often used for test validation, internal QC, method development, and assay calibration, the survey showed. Dr. Barton believes that commonly available reference materials are most often used for daily run controls. “Certified reference materials, the top-level material available from NIST and WHO, tend to be quite expensive,” he says. “I expect labs would use those only occasionally to calibrate assays.” Dr. Barton uses reference materials across the board in his laboratory. “We’re a national centre,” he points out, “and I’ve spent many years advocating their use.”

Another group with which Dr. Barton works is the European Molecular Quality Network, or EMQN. “We are the biggest provider of EQA or PT for genetic disorders in the world,” he says. A unique feature of EMQN is that it assesses not just genotyping accuracy but also the interpretation of genotype. “We set clinical cases for every QA,” Dr. Barton says. “Our evaluators assess whether the genotype is correct plus whether the interpretation is correctly given for that case.” Last July Dr. Barton spoke to the CAP’s Next-Generation Sequencing Working Group, which is working on standards development, proficiency testing, and other NGS-related issues. He emphasized accuracy of interpretation. “We think that’s fundamentally important,” he told CAP TODAY. EMQN offers proficiency tests for 25 hereditary disorders, such as familial breast cancer, hereditary deafness, and fragile X syndrome.

While expensive certified reference materials won’t be used in daily routine, assays can be calibrated to a high level of accuracy by using the certified materials to calibrate EQA materials, which can then be used to calibrate laboratory controls. In a 2012 PT challenge for Huntington disease diagnosis, EMQN gave laboratories an opportunity to adjust their Huntington disease PT material sizing to the NIST standard. Genotype (triplet repeat numbers) of the PT material was verified with a NIST standard reference material. Participants could then calibrate their assays using the PT materials.

Counting GAA repeats in Friedreich ataxia presents a similar conundrum. In a PT challenge, laboratory results showed substantial scatter. “The range wouldn’t change the category,” Dr. Barton says. “All alleles reported as disease-causing were disease-causing, and those reported as in normal range were in the normal range. However, the scatter raises a concern that, if you did have an allele close to the borderline value, some labs would get it wrong.”

“All labs think they are reporting correct allele size,” he adds. “Only through EQA or PT can you show people they are not conforming with the consensus.” For Huntington disease this is easier, since a NIST standard value can be provided, which is powerful in the argument. “Otherwise,” Dr. Barton says, “strong-minded lab directors will say, ‘I’m right, they’re wrong.’” Well-characterized standards for Friedreich ataxia, which are lacking, would serve the same function.

Genomic reference materials are also available from the National Institute for Biological Standards and Control (NIBSC), a government-funded body much like NIST, which, Dr. Barton says, is “the only laboratory in the world making biological reference materials for WHO certification, not just in genetics but across the scope of clinical laboratory testing.” NIBSC offers 11 genomic reference materials at this time, and three more are in development (www.nibsc.ac.uk).

CRMGEN, a preview project that Dr. Barton chaired, determined what type of reference materials would be best for genetic testing. “What people want most,” he says, “the most versatile format in everyday life, is genomic DNA.

Anything else restricts the usefulness of reference material to specific assays.”

Highlighting the necessity to use well-characterized reference materials, Dr. Barton cites detection of the R117H mutation in cystic fibrosis. “In most populations this mutation is moderately rare. So homozygous mutant individuals are not seen.” In Ireland, on the other hand, there is a high frequency of R117H. “So we have several samples from patients homozygous for this mutation. When we shared that material with test developers, on two occasions they found their normal signal was not discriminating between normal and mutant alleles. It was not specific for the normal allele. So they got a heterozygous signal on our homozygous mutant sample.”

Turning to the U.S. experience, Dr. Barton says he is an admirer of Dr. Kalman’s work at the CDC. “That program has taken a very pragmatic approach to characterizing existing materials. Making certified materials is an extremely expensive business. If they had set out to do that, they might have two or three certified reference materials on the market. Instead, they have many dozens of verified control DNAs that may not meet the standard of CRMs but are still useful to labs.”

Dr. Kalman coordinates the Genetic Testing Reference Materials Program (GeT-RM), which evolved from a pilot program completed in 2003. After that effort successfully created and characterized 27 new cell lines, the GeT-RM was initiated in 2004. Since that time, the GeT-RM has characterized more than 300 cell lines. It is essentially an ad hoc program. “People become involved as we do different projects,” Dr. Kalman, health scientist in the CDC Division of Laboratory Science and Standards, said in an interview. Volunteer laboratories characterize the reference materials. “When you think about this, it can cost thousands of dollars for them to run these samples,” she says. “And we do not reimburse them. They do it because they feel these materials are necessary.”

Reference materials can be of varying quality, Dr. Kalman says, with the degree of characterization being the chief variable. “It is well defined as to what you have to do to meet the requirement for a certified or standard reference material,” she says. QC materials, which include genomic DNA, are homogeneous and stable but not necessarily well characterized. What certified and standard reference materials have in addition is a certified value, along with its uncertainty, and stated traceability to a known standard, original standard, or reference method. Dr. Kalman estimates that about 20 standard reference materials, certified reference materials, or WHO standards are available, along with about 300 characterized genomic DNAs resulting from the work of GeT-RM. GeT-RM has created characterized genomic DNA reference materials for a number of disorders, including fragile X, Huntington disease, BRCA1/2, and Duchenne muscular dystrophy. In development are reference materials for Rett syndrome, pharmacogenetics, and cytogenetics.

“We do exclusively genomic DNA,” Dr. Kalman says. All of the genomic DNAs characterized by GeT-RM come from Coriell repository cell lines, which originated from patients with a particular disease. To emphasize the meaning of genomic DNA, Dr. Kalman says, “NIST’s fragile X standard is a PCR amplicon, not genomic DNA. It is a ‘synthetic DNA molecule’ amplified from a human cell line.” She points out, as does Dr. Barton, that genomic DNA from cell lines doesn’t have the same function as certified or NIST reference materials. GeT-RM cell line DNA is logistically and economically feasible to use for daily controls, she says.

Developing reference materials for Duchenne muscular dystrophy is a situation in which GeT-RM had to develop new cell lines for some mutations, since existing cell lines covered only deletions, not point mutations or duplications in the *DMD* gene. Working from a patient registry with known mutations sponsored by a patient advocacy group, and going through institutional review board approval and patient consent, researchers collected blood from patients with mutations of interest, and Coriell made 10 new cell lines. Both male probands and female carriers were represented. “We’ve facilitated the creation of new cell lines for a number of projects,” Dr. Kalman says, including the ongoing Rett syndrome work.

New projects for GeT-RM include pharmacogenomics, cytogenetics, molecular oncology (with Dr. Jennings), and next-generation sequencing. “We are now working on reference material for the whole genome sequence,” she says.

For the NGS project, existing and new sequence data are being collected for two human cell lines from more than

36 clinical gene panels, exome and whole genome tests. As usual, volunteer laboratories are doing the work. The resulting format will be not one sequence but what each laboratory produced with its method. Information collected includes an assessment of data quality, such as coverage and quality scores. Data are sent to the National Center for Biotechnology Information, which is building a whole genome browser to allow access to the reference material sequence data in collaboration with clinical laboratories.



Dr. Kalman

Dr. Kalman distinguishes between these reference genomes, which will represent two specific individuals with determined sequences, and the sequence from the Human Genome Project, which averaged data at each base position from many individuals. The purpose of the NGS project is not for diagnosis, since these are two supposedly healthy individuals. “We are trying to characterize the sequence of these samples,” Dr. Kalman says, “so that when someone wants to validate a new next-generation sequencing assay, they can buy samples from Coriell and compare their sequence to our results. That should enable them to troubleshoot their assay.”

New challenges will arise from the NGS project. “The quantity of data will be enormous,” Dr. Kalman says. Also, it is difficult to characterize a genomic sequence with respect to SI units, as one can do with clinical chemistry analytes, such as sodium or glucose. These chemicals can be precisely quantified, but not so for genomic DNA. “NIST is currently creating highly characterized reference materials for the whole human genome, three billion base pairs with four possible nucleotides in each position,” Dr. Kalman says. “The result is not quantitative, not how much of each nucleotide, just a sequence.” Each sample will contain three billion analytes. “Even one gene is orders of magnitude more difficult than anything we have done previously,” she says. “We will have to develop a different model of how you make this new reference material, of how dependable it has to be to make a medical genetic diagnosis.” Even with these differences, the NGS project will adhere to the same underlying principle that guided previous projects: Truth in laboratory testing requires well-characterized reference materials.□

William Check is a writer in Ft. Lauderdale, Fla.