Medical genetics labs shine in 10-year proficiency test data

William Check, PhD

January 2015—Molecular genetics laboratories in the U.S. are doing a great job.

Ten-year data from the molecular genetics Surveys in the CAP proficiency testing program show that U.S. clinical laboratories are making extremely accurate calls using molecular genetics assays.

At the 2014 meeting of the Association for Molecular Pathology, Karen E. Weck, MD, and Iris Schrijver, MD, presented results from seven of the proficiency Surveys that the CAP/ACMG Biochemical and Molecular Genetics Resource Committee oversees. Dr. Weck is the chair of the committee; Dr. Schrijver is past chair.

“We have seen excellent performance of molecular genetics testing in general,” Dr. Weck says. “Overall sensitivity has been greater than 95 percent and specificity greater than 99 percent for all analytes we have evaluated thus far.” (Analysis of females for Fragile X had a sensitivity of 95.5 percent; all other sensitivities were around 99 percent.)

“Laboratories are doing a good job across the board,” Dr. Schrijver tells CAP TODAY. The committee members selected seven Surveys for analysis ranging from commonly performed cystic fibrosis testing to complex Fragile X syndrome testing, which has multiple components. They published results from these analyses over the past two years in a series of articles in *Genetics in Medicine*.

Analyzing molecular genetics laboratory performance is important for two reasons, says Dr. Weck, director of molecular genetics and professor of pathology and laboratory medicine and of genetics at the University of North Carolina. First, it can inform the molecular genetics laboratory community what expectations should be in terms of different methods. Second, it helps to identify areas where efforts need to be made to improve testing, “such as standardization or calibration or guidelines for interpretation,” she says.

Surveys provide more than numbers, says Dr. Schrijver, director of the molecular pathology laboratory at Stanford University Medical Center and professor of pathology and of pediatrics, Stanford University School of Medicine. “It is important not just to look at the data in the Surveys, but to provide educational content. Laboratories have improved considerably over time, particularly in the interpretive analysis,” Dr. Schrijver says. “They’re learning from participating in the Surveys.”

Further, results of these analyses shed light on the performance of laboratory-developed tests. “Of the greater than 30 molecular genetics Surveys overseen by our committee, almost all are performed by LDTs,” Dr. Weck points out. All of the methods used in the seven Surveys analyzed were laboratory-developed tests. “There are no U.S. Food and Drug Administration–approved test platforms for most of these disorders,” Dr. Schrijver says. The data show molecular genetics laboratories to be highly accurate when using LDTs.

In addition to the CAP analyte-specific molecular genetics Surveys, methods-based PT for Sanger sequencing has been offered for several years, which provides a segue to PT for next-generation sequencing. The CAP is launching a methods-based PT program for NGS, with the first PT mailing to take place in March. “Proficiency testing for next-generation sequencing will definitely be more challenging,” Dr. Weck says.

Dr. Voelkerding

Karl V. Voelkerding, MD, is chair of the CAP Next-Generation Sequencing Project Team, which is responsible for developing laboratory accreditation requirements and proficiency testing for next-generation sequencing. “In 2013–2014 we conducted a pilot program to assess a potential proficiency testing program for next-generation sequencing,” says Dr. Voelkerding, professor of pathology at the University of Utah and medical director for genomics and bioinformatics at ARUP Laboratories. “The results of that pilot were sufficiently encouraging to move forward and launch an educational PT program in 2015. When we receive and review results from participating laboratories, they will be used to refine the proficiency testing program.” The first next-generation sequencing-based PT program will focus on the ability of laboratories to detect sequence variants in germline DNA.

Dr. Schrijver credits the late Jeffrey Kant, MD, PhD, for the decision to analyze Surveys in molecular genetics. “This was a vision that originated with Jeff Kant to document performance of laboratories on individual assays and to share that information publicly,” she says. When Dr. Schrijver was chair of the CAP/ACMG committee, she put that plan in motion. “We were not sure what we would find,” she says. “We went back as far as each Survey existed.” CAP staff did the basic analysis, first-level interpretations, and summaries for the committee. All data were anonymized.

There was no selection bias among laboratories, Dr. Schrijver says. “All diagnostic laboratories in the U.S. have to be CLIA-certified and accredited by some organization. Virtually all molecular genetic laboratories are accredited by the CAP.”

For the Fragile X syndrome Survey, DNA from Coriell cell lines with known FMR1 genotypes was distributed. Individual laboratory responses were analyzed for accuracy of genotype determination and triplet repeat size. Sizing of the CGG triplet repeat region can be extremely important in some size ranges.

Sensitivities for full mutations were 99 percent in males and 95.5 percent in females; sensitivity for premutation was 98 percent. Specificity was 99.9 percent (Weck KE, et al. *Genet Med.* 2012;14:306–312). “Interlaboratory sizing by PCR improved over time, coincident with an increase in use of capillary electrophoresis over gel-based PCR and the availability of well-characterized materials for calibration,” Dr. Weck said.

For Huntington disease, diagnosis depends on detection of CAG repeat size in the HTT gene. By this criterion, analytic sensitivity found in the Survey was 99.5 percent and analytic specificity was 99.2 percent (Palomaki GE, et al. *Genet Med.* 2012;14:69–75). Plotting performance of individual laboratories over time showed that U.S. laboratories performed significantly better than non-U.S. laboratories. (This was observed in a few other Surveys as well.) “International laboratories are not CLIA-certified,” Dr. Schrijver notes. “They might be research laboratories.”

Sensitivity and specificity for Ashkenazi Jewish founder mutations in the BRCA1 and BRCA2 genes were found to be 99 percent and 99.9 percent, respectively (Tafe LJ, et al. *Genet Med.* Published online June 19, 2014. doi:10.1038/gim.2014.77). Only an average of 23 U.S. laboratories participated per year in the Survey, most likely due to patent issues. “Now that the patent has lifted, we will probably see more laboratories participating,” Dr. Weck says. Another consequence of the patent lifting is that the committee plans to extend the Survey to full gene-sequence analysis. Even so, the three mutations that have historically been analyzed by the Survey account for 95 percent of pathogenic mutations in the Ashkenazi Jewish population.

Dr. Weck

While analytical accuracy was very high, clinical interpretation was slightly less accurate. “Most interpretive errors—81 percent—were due to incorrect interpretation of the lifetime risk of breast cancer for a patient with a negative result,” Dr. Weck says. The proficiency testing samples were identified as from a woman with an Ashkenazi Jewish background and a strong family history of early-onset breast cancer. Participants were given a multiple-choice question, with the best answer for a negative result being “The lifetime risk of breast cancer cannot be determined without BRCA mutation testing of the affected relative.” The correct response was given in 92.5 percent of cases. “The woman is still at greater risk than the average population lifetime risk based on her family history,” Dr. Weck explains. Without knowing whether the affected relative has a BRCA mutation, the negative result is uninformative. For instance, a deletion would not be picked up by Sanger sequencing. (Three-fourths of participants used Sanger sequencing.) However, Dr. Weck cautions, “Our ability to analyze how laboratories interpret their own results in their laboratory reports is somewhat limited.”

The cystic fibrosis Surveys provided from 2003 to 2013 were designed to test performance of the ACMG/ACOG recommended 23-gene panel. One hundred seventy-nine laboratories participated. Overall analytical sensitivity was 98.8 percent and specificity was 99.6 percent (Lyon E, et al. *Genet Med.* Published online July 31, 2014. doi:10.1038/gim. 2014.93). Dr. Schrijver says there was a significant trend toward higher analytical sensitivity between 2003 and 2008. Only 21 percent of Surveys in which an error occurred were distributed after 2007.

However, as with the BRCA Survey, there were interpretive errors, with the greatest variation seen in how laboratories interpreted the clinical significance of heterozygous samples. In the Survey, the clinical scenario was a child with failure to thrive. In this context, there is a spectrum of possible findings, including two distinct mutations. “One thing we have been educating around is that laboratories should know that many more mutations are possible in people with CF than the 23 on the panel,” Dr. Schrijver explains. “So finding one mutation on the panel doesn’t mean that CF is less likely.” In 10.8 percent of cases where one mutation was found, the laboratory gave this incorrect interpretation.

A Survey for genotyping and interpreting the basic Ashkenazi Jewish panel—with multiple conditions in addition to Tay-Sachs disease, Canavan disease, and familial dysautonomia—has been provided since 2006. Analytic sensitivity was 97.2 percent and specificity was 99.8 percent; analytic interpretations were correct in 99.3 percent of challenges (Feldman GL, et al. *Genet Med.* 2014;16:695–702). Laboratories provided accurate test results in both diagnostic and screening settings.

The number of people tested with the Ashkenazi Jewish panel increased in 2011 about fivefold, from approximately 2,500 per month to around 12,500 per month. “At first we thought there could be multiple reasons,” Dr. Schrijver recalls. “Now we know that a few laboratories started offering large carrier screening panels to anyone, regardless of ethnicity.” Expanding the pool of people screened to those outside the Ashkenazi Jewish community can be a problem—those with different ethnic backgrounds have different mutations not captured by typical panels. So a negative finding could give people a false sense of security. Dr. Schrijver acknowledges that the laboratories that are screening non-Jewish persons may have expanded their panels to contain a large number of mutations. “But we don’t know that,” she says.

CAP proficiency testing for Ashkenazi Jewish conditions was expanded in 2012 to include Gaucher disease, Bloom syndrome, Niemann–Pick disease type A, glycogen storage disease type 1a, mucolipidosis type IV, and Fanconi anemia type C.

Of the overall high quality of the results on the molecular genetics Surveys, Dr. Weck says: “Expertise in molecular

testing by clinical laboratories is at a high level. Molecular pathologists and directors of molecular laboratories are molecular professionals who apply rigor in their own work and are very nimble at being able to take a new technology or technique, validate it very quickly, and use it to detect whatever new analyte needs to be tested."

Moving from analyte-specific proficiency testing to methods-based testing, Dr. Schrijver noted that this approach is already established in CAP PT programs for immunohistochemistry, FISH, and flow cytometry. "Now with the advent of next-generation sequencing, which has a very large scope of genetic testing, analyte-specific proficiency testing is not feasible anymore," she says.

"There are about 22,000 genes that could be potentially tested in whole exome sequencing; that is what you are going for. There is no way the College could provide proficiency testing for all of these genes. Yet you want to make sure that laboratories that do larger-scope testing can identify and name mutations correctly and interpret whether they are pathogenic or benign" (Schrijver I, et al. *J Mol Diagn.* 2014;16:283–287).

The Sanger sequencing PT program has the advantage of covering testing for rare diseases. These are conditions for which there are no interlaboratory exchange partners. When an analyte-specific test is available, Dr. Schrijver says, the laboratory has to subscribe to it also. "That's because gene-specific expertise can be measured and you can address more detailed interpretation of individual mutations, which is different from methods-based testing."

Dr. Schrijver

In the sequencing educational challenge (SEC), a dry test for Sanger sequencing, "the laboratory knows which gene is being sent, which is not necessarily a gene they test for in their lab," Dr. Schrijver says. "CAP sends three challenge files with sequences and three normal ones and asks the laboratory to interpret the data." Challenges require the laboratory to address zygosity and use correct Human Genome Variation Society (HGVS) nomenclature, to identify all variants and say whether they would change the protein, and to provide a basic interpretation, adding up to four questions per abnormal file, for a total of 12 answers. "A passing grade is 10 of 12 correct," Dr. Schrijver explains. Of the 67 U.S. participants, 98.3 percent had acceptable performance, compared with 88.9 percent for the 50 international participants (Richards CS, et al. *Genet Med.* 2014;16:25–32). "These data provide a high level of confidence that most U.S. laboratories offering rare disease testing are providing consistent and reliable clinical interpretations," the authors concluded.

Areas in which there is room for improvement include correctly naming predicted proteins for frameshift mutations, following HGVS nomenclature rules, and following guidelines for interpretation of pathogenicity.

A wet challenge became available in 2013 for Sanger sequencing. Laboratories receive three DNA samples plus primers and are asked to generate a sequence and interpret it.

Dr. Schrijver foresees more methods-based testing in the near future, such as for multiplex ligation-dependent probe amplification and chromosomal microarrays.

"Now it is time to apply this approach to next-generation sequencing," she says. About two-thirds of participants in the molecular genetics Surveys expected to introduce NGS technology in the near future. "We realized that proficiency testing for next-generation sequencing is a logical and important next step to ensure correct variant identification," Dr. Schrijver says. But applying proficiency testing to NGS is an entirely different level of complexity. For whole genome sequencing, she notes, the number of variants per person is about 3 million; for whole exome sequencing it is between 15,000 and 20,000.

The genomic DNA used in the NGS pilot PT program, Dr. Voelkerding says, was sourced from an individual who gave extensive informed consent. “This genomic DNA was subjected to exome and whole genome sequencing,” he says, “followed by bioinformatics analysis to derive a consensus set of sequence variants and wild-type reference positions.”

In the pilot, genomic DNA was sent to eight laboratories. They were asked to query up to 200 genomic loci consisting of a mix of single nucleotide variants, indels, and wild-type nucleotides. Some laboratories did targeted gene panels; others did exome sequencing. Three laboratories had 100 percent correct identification of the genomic loci they assessed. Three others were between 91 percent and 97 percent. For the two laboratories with lower percentages, Dr. Voelkerding says, “We think there were typographical errors on the part of the reporting laboratory. That allowed us to revisit how to structure the test result form to minimize incorrect answers due to typographical errors.”

With results from the pilot program considered satisfactory, an educational Survey will begin in March. “Operationally, that means that proficiency testing will be administered twice in 2015 and results will be returned to laboratories so they have an understanding of how they performed, but they will not be officially graded,” Dr. Voelkerding says. “If proficiency testing as designed appears appropriate with acceptable results, I would anticipate in 2016 we would move to formal proficiency testing, where laboratories that sign up and participate would not only receive results but would receive a grade in relationship to all laboratories.”

As an indication of how quickly laboratories are adopting NGS, and how urgently PT for NGS is needed, Dr. Voelkerding says about 130 laboratories recently indicated the activity code for NGS on their test menu update submitted to the CAP. “What is interesting about that number,” he says, “is that a couple of years ago it was more like 25. So the number of laboratories listing the activity code for NGS has gone up fivefold over a couple of years.”

Some laboratories are using NGS for germline disorders or inherited disorders and others to detect somatic variants and mutations in cancer biopsies. “Because there are differences between those two applications,” Dr.Voelkerding says, “new NGS proficiency testing for somatic mutations is also being developed, and some of the available PT challenges developed by the CAP Molecular Oncology Committee can be used for NGS PT assessment for specific somatic mutations.”

Dr. Weck calls the design of the proficiency test for germline mutations “pretty slick.” Laboratories report variants for particular variant positions and genomic coordinates that are given to them. “And they only report for genes for which they are doing clinical testing,” she explains. “So we can offer a broad Survey but allow laboratories to target their responses to those genes for which they have expertise.” It will be a challenge to the CAP to analyze and collate all the results, she adds. “It will be very informative in the educational phase to see how much data CAP gets back and how easy it will be to give results to participating laboratories.”

Developing proficiency testing that is robust and accurately assesses laboratories’ proficiency takes time, Dr. Voelkerding notes. It will be about three years from conception of the pilot to the educational phase to launching a formal graded PT. “Within the College we are discussing how to meet the demand to create a robust proficiency testing process as expeditiously as possible.” Much has been learned in developing the first NGS PT, he says. “Subsequent Surveys should be accelerated just by virtue of our having undergone a learning curve already.”

“What I envision,” he continues, “is essentially a portfolio of proficiency testing that will encompass traditional applications of Sanger sequencing as well as new approaches and applications of next-generation sequencing in areas of inherited disorders and molecular oncology, both solid tumors and hematologic malignancies.” Next-generation sequencing is now in the developmental phase for infectious diseases and for tissue typing with HLA markers. “Sanger sequencing is employed for tissue typing, but there are technological and logistic and cost-efficiency gains that can be made by introducing next-generation sequencing into HLA testing,” he says.

Proficiency testing addressing all those applications of NGS has to be developed. “Our overarching challenge,” Dr. Voelkerding says, “is sourcing appropriate materials for proficiency testing—the specific samples that will be

used—and designing Surveys to accurately gauge the ability of laboratories.” The job ahead is big indeed. But with labs using the traditional methods for molecular genetics testing having set a high performance bar, the motivation is strong to ensure the same for next-gen sequencing.
[hr]

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