

Results from an external proficiency testing program: 11 years of molecular genetics testing for myotonic dystrophy type 1

C. Sue Richards, PhD¹, Glenn E. Palomaki, PhD² and Madhuri Hegde, PhD³; on behalf of the Biochemical and Molecular Genetics and Genomics Resource Committee, College of American Pathologists/American College of Medical Genetics and Genomics

Purpose: The aim of this study was to examine the performance of laboratories offering assessment for myotonic dystrophy type 1 (DM1) using external proficiency testing samples. DM1, a dominant disorder, has a prevalence of 1:20,000 due to the expansion of CTG trinucleotide repeats in the *DMPK* gene.

Methods: External proficiency testing administered by the College of American Pathologists/American College of Medical Genetics and Genomics distributes three samples twice yearly. Responses from 2003 through the first distribution of 2013 were analyzed after stratification by location (United States/international). Both the repeat sizes (analytic validity) and clinical interpretations were assessed.

Results: Over the 21 distributions, 45 US and 29 international laboratories participated. Analytic sensitivity for detecting and reporting expanded repeats (≥ 50) was 99.2% (382/385 challenges) and 97.1%

(133/137 challenges), respectively. Analytic specificity (to within two repeats of the consensus) was 99.2% (1,790/1,805 alleles) and 98.6% (702/712 alleles), respectively. Clinical interpretations were correct for 99.3% (450/453) and 98.2% (224/228) of positive challenges and in 99.9% (936/937) and 99.6% (455/457) of negative challenges, respectively. Of four incorrect interpretations made in the United States, two were probably due to sample mix-up.

Conclusion: This review of laboratory performance regarding laboratory-developed genetic tests indicates very high performance for both the analytic and interpretative challenges for DM1.

Genet Med advance online publication 2 June 2016

Key Words: external proficiency testing; myotonic dystrophy type 1; trinucleotide repeats

INTRODUCTION

Myotonic dystrophy type 1 (DM1) is an autosomal dominant disorder caused by mutations in the *DMPK* gene located on chromosome 19q13.32 (ref. 1) that affects the muscle, eye, heart, endocrine, and central nervous systems. The overall worldwide prevalence of this disease is approximately 1:20,000 (ref. 2). The disease results exclusively from the expansion of a CTG trinucleotide repeat, located in the 3' noncoding region of the *DMPK* gene. Normal-sized alleles contain 5–34 CTG repeats, whereas premutation alleles contain 35–49 repeats and are not associated with symptoms but can expand in future generations. Full expansion alleles comprise ≥ 50 CTG repeats and are associated with disease with variable expressivity.^{3–6} DM1 is a classic example of anticipation in families. Longer CTG repeats correlate with greater disease severity and earlier onset, and penetrance is nearly 100% by age 50 (ref. 4). The clinical categories of DM1 include (i) mild disease with 50–150 CTG repeats that is characterized by cataract and mild myotonia with a normal life span; (ii) classic disease with 100–1,000 repeats that is characterized by muscle weakness and wasting, cataract, and abnormal cardiac conduction with disability and a reduced life

span; and (iii) congenital disease with more than 1,000 repeats that is characterized by hypotonia and severe weakness at birth, respiratory problems, early death, and intellectual disability.^{3,4,7} The abnormal gene product is a gain-of-function RNA that interferes with RNA transcript processing. Some proteins affected are directly related to symptoms of DM1, including myotonia (chloride channel protein), increased risk of diabetes mellitus (insulin receptor protein), and cognitive function (microtubule-associated protein tau).^{7–15} Both diagnostic testing in symptomatic individuals and family members at risk and prenatal testing are routinely performed in molecular genetics laboratories through analysis of CTG repeat length.

DM1 is characterized by large trinucleotide repeat expansions and thus requires both polymerase chain reaction to determine the size of the normal allele and the premutation allele and Southern blotting to determine the size of the larger expansions. In the past, having sufficient standards for developing the allelic ladder has been a challenge. Precise measurements are critical, especially when results occur near the boundary between size categories. Test requests for fragile X, another trinucleotide repeat disorder, are far more common, and several US Food

¹Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon, USA; ²Department of Pathology and Laboratory Medicine, Women & Infants Hospital and the Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA; ³Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA. Correspondence: C. Sue Richards (richarsu@ohsu.edu)

Submitted 2 November 2015; accepted 28 March 2016; advance online publication 2 June 2016. doi:10.1038/gim.2016.59

and Drug Administration (FDA)-approved test reagents are commercially available. Because of the low number of DM1 test requests, however, no FDA-approved reagents are currently available. Therefore, DM1 testing in the United States requires the use of laboratory-developed tests whereby the laboratory is responsible for developing a quality-control system to ensure the accuracy of results. Because of the importance of reliable testing for DM1 alleles and their accurate interpretation, the College of American Pathologists (CAP) offers external proficiency testing (PT) challenges for DM1 molecular genetic testing twice per year. The current analyses cover an 11-year period, from 2003 through the first half of 2013. Specific questions are addressed: (i) Among current participants, how many clinical tests are performed and what is the turnaround time? (ii) What is the analytic performance of laboratories located in the United States? (iii) Is the performance of international laboratories similar? (iv) How reliable are the associated interpretations, after stratification by laboratory location?

MATERIALS AND METHODS

For each of two annual Molecular Genetics Surveys (MGL), participating molecular diagnostic laboratories receive three DNA sample challenges selected from validated reference materials with known genotypes (Coriell Institute for Medical Research, Camden, NJ). Laboratories use their own proprietary, internally validated tests. The results of the PT challenges are submitted to CAP, where they are analyzed, reviewed, and summarized, then forwarded to the joint CAP/American College of Medical Genetics and Genomics (ACMG) Biochemical and Molecular Genetics Resource Committee for review and final interpretation. Assessment of laboratory performance is based on both the genotype and its clinical interpretation. Results are confidential between CAP and the laboratory, but laboratories can disclose their performance on PT programs upon request. If a laboratory fails the DM1 PT challenge, they must immediately address relevant issues in communication with CAP. If a laboratory is accredited by the CAP Laboratory Accreditation Program and has unsatisfactory performance on PT, the laboratory receives a notice regarding the failure. If there are multiple failures for the same analyte, the laboratory is required to provide corrective action or could be directed to cease testing for that analyte. The mission of the Laboratory Accreditation Program is to advance the quality of laboratory services through PT, education, and unannounced on-site inspections by trained inspectors to ensure that regulatory requirements are being met under the authority granted by CLIA. The committee members have no direct association with Laboratory Accreditation Program activities and do not contact individual laboratories to determine the root cause of incorrect responses.

Analytic specificity challenges were those with a consensus of <35 CTG repeats for both alleles. Any reported CTG repeat size within two of the consensus (median) was considered correct. If a laboratory reported only one allele (e.g., “17, blank”), it was assumed that the laboratory was reporting a homozygous result. For example, the response “5, blank” was considered

incorrect for a specificity challenge with two different negative alleles (e.g., 5, 13) but was accepted as correct for a homozygous challenge (e.g., 5, 5). Because both alleles were graded, the potential number of responses is twice the number of “normal” sample challenges.

Analytic sensitivity challenges were those having a consensus of 50 or more CTG repeats on one allele. All sensitivity challenges involved a mosaic pattern with well over 150 repeats. Since this is an autosomal dominant condition, the second allele is usually of normal length but is not included in the assessment of analytic sensitivity. Thus the potential number of responses is equal to the number of “abnormal” sample challenges. An incorrect sensitivity response was noted if (i) the reported abnormal repeat length was <50 or (ii) the repeat length was not specified (e.g., blank) but the clinical interpretation was “negative” or was not reported. This second definition was needed because some laboratories observed an allele with a large repeat size on Southern blotting, but they do not quantify the number of repeats for clinical reporting. This occurs in spite of CAP and other relevant professional recommendations (ACMG) to quantify large DM1 repeat lengths. Thus, the response “5, blank” was considered correct if the laboratory reported a “positive” clinical interpretation but incorrect if it was associated with a “negative” clinical interpretation.

The associated clinical interpretations have changed over time, but they can be summarized as “negative—not consistent with DM1,” “mutable allele, normal” and “positive—consistent with DM1.” These three interpretations are associated with repeat CTG sizes of <35, 36–49, and ≥50, respectively. Among those with 50 or more repeats, it is possible to provide some phenotypic information.^{2,5} No samples with “mutable allele, normal” have been distributed as part of this PT program. If a laboratory reports an incorrect analytic result, it is often associated with an incorrect clinical interpretation. Current grading does not consider whether the clinical interpretation would have been correct, given the incorrect analytic result reported.

Confidence intervals for proportions were at the 95% level (two-tailed) (True Epistat, Round Rock, TX) using a binomial distribution. Statistical tests of significance for contingency tables were exact two-tailed, with $P < 0.05$ indicating significance.

RESULTS

Between 2003 and 2013, 45 US and 29 international laboratories participated in at least one DM1 PT survey. The numbers of participants increased over time, from 24 in 2003-A (22 from the United States) to 40 in 2013-A (22 from the United States). The primary testing method was almost always based on polymerase chain reaction, whereas the secondary method (for large repeat lengths) was Southern blotting. One laboratory reported using DNA sequencing. As part of each distribution, participants were asked to report the number of tests performed and clinical testing turnaround time in their laboratory. Data from 2012-A and 2013-A were used to compare the number of tests

versus turnaround time, shown in **Figure 1**, which contains data from 25 US and 19 international participants. Overall, an estimated 5,376 clinical tests were performed annually (3,360 in the United States and 2,016 internationally). Among US laboratories, 64% (16/25) performed fewer than five tests per month, whereas that rate among international laboratories was 37% (7/19; $P = 0.13$). The US laboratories tended to report more quickly, with 88% (22/25) posting results within 21 days compared with 37% (7/19; $P < 0.001$) of international participants posting results within that time. Given the observed digit preference, at least some of the reported turnaround times are likely to be quoted rates and not actual measured times.

The DM1 genotyping challenges and a summary of participant results used to compute the analytic sensitivity and specificity are provided in **Supplementary Table S1** online. Overall, the US laboratories had an analytic sensitivity of 99.2% (382/385 challenges; 95% CI: 97.7–99.8%), which is not significantly different from the 97.1% (133/137 challenges; 95% CI: 92.7–99.2%) for international laboratories ($P = 0.08$). Analytic specificity was also high at 99.2% (1,790/1,805 alleles; 95% CI: 98.6–99.5%) and 98.6% (702/712 alleles; 95% CI: 97.4–99.3%), respectively. These are also not significantly different ($P = 0.19$). Of the 45 US laboratories, 10 (22%) had analytic errors (**Table 1**), and 6 of these still participate in the survey. Of the 29 international laboratories, 7 (24%) had analytic errors, and 3 of these still participate. Of the 74 laboratories, 9% (7/74) have made multiple errors, but these 7 laboratories are responsible for 71% (24/34) of all analytic errors. The total analytical error rate for the 27 US laboratories testing fewer than five samples per month was 1.1% (95% CI: 0.5–2.1%), which is not

significantly different from the 1.9% (0.9–3.5%) for the 16 laboratories processing five or more samples per month ($P = 0.25$).

Although the analytic sensitivity challenges were mosaic for DM1 repeat lengths, specific instructions were given by the CAP to report the major (densest) band rather than the band(s) corresponding to the smallest or largest numbers of repeats. One positive sample was distributed six times over the 11 years, with median consensus values of 600, 670, 630, 660, 670, and 660 CTG repeats. We compared the reported within-laboratory results to determine whether there was a consistent reporting protocol being used that might differ from that recommended or that used in other laboratories. **Figure 2** shows the results of comparing the two of the six distributions where the consensus results were nearly identical at 650 repeats. The open circles indicate the 22 laboratories reporting results for both challenges. Although the majority of the paired results are close to the consensus results, reflecting excellent long-term precision, there are several clear discrepancies. For example, two

Table 1 Distribution of analytic errors by laboratory location

Errors (n)	US laboratories, n (%)	International laboratories, n (%)
0	35 (78)	22 (76)
1	6 (13)	4 (14)
2	2 (4)	0 (0)
3	1 (2)	2 (7)
4	0 (0)	1 (3)
5	1 (2)	0 (0)

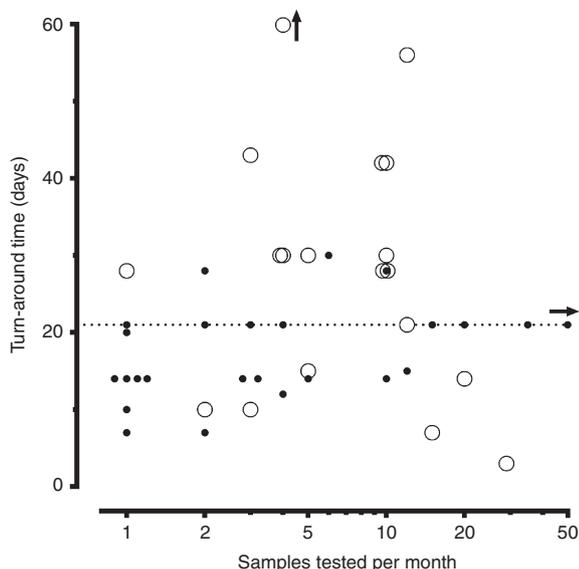


Figure 1 Numbers of samples tested per month versus reported turnaround time for participants in an external proficiency testing program for myotonic dystrophy type 1 (DM1), stratified by laboratory location. Smaller filled circles represent the 25 laboratories located in the United States; larger open circles represent the 19 international laboratories. The horizontal dotted line is drawn at a turnaround time of 3 weeks (21 days). The small arrows indicate values greater than the axis upper limit.

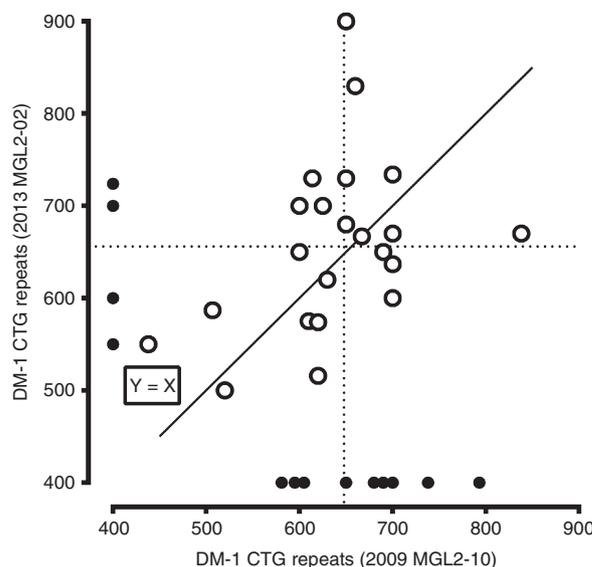


Figure 2 Reported CTG triplet repeat length for two identical sample challenges distributed 4 years apart. Open circles indicate results for the 22 participants reporting repeat length for both challenges. Most results are close to the consensus of 660 repeats for both challenges; however, some outlying values occurred. The filled circles show results for laboratories participating in only one or the other sample challenge.

laboratories reported almost exactly the consensus 650 repeats in 2009 (x axis) but reported 830 and 900 CTG repeats for the same sample in 2013 (y axis). This may be due to changes in personnel and/or differences in reporting mosaics, or to assay changes related to sample digestion, electrophoresis, and blotting. Filled circles along the x or y axis indicate laboratory results for those participating in only one or the other distribution.

In addition to genotyping, clinical interpretations were reported by most participants for each challenge. These data were also analyzed for sensitivity and specificity after stratification by laboratory location. There tended to be fewer clinical responses (compared with analytic responses) for positive sample challenges because some laboratories (especially those outside the United States) did not quantify/report sizes when the repeat lengths were over 50. Among US laboratories, clinical interpretations were correctly reported as “positive, consistent with DM1” for 99.3% (450/453; 95% CI: 98.1–99.9%) for samples with consensus large repeat lengths. These three false-negative errors can be seen in **Table 1** (2008B-10, 2011A-3 and 2012B). In the first two, the laboratories did not identify the expanded allele and reported a negative clinical interpretation. However, all other respondents identified an allele with 620, 400, and 350 repeats, respectively. In the 2012B distribution, there was a likely sample mix-up in one US laboratory. The three correct genotypes and associated interpretations were reported, but not for the appropriate challenge. Among international laboratories, there were four errors among positive sample challenges (2004A-03, two in 2008B-10, and 2010A-01; **Table 1**). The corresponding rate is 98.2% (224/228; 95% CI: 95.6–99.5%). All errors involved laboratories that did not identify the expanded allele (e.g., analytic error causing the error in clinical interpretation). One lab was involved in two of these errors. The rates of correct clinical interpretation of “positive, consistent with DM1” do not differ between US and international laboratories ($P = 0.23$).

Among US laboratories, clinical interpretations were correctly reported as “negative—not consistent with DM1” for 99.9% (935/936; 95% CI: 99.4–99.9%) of samples with both alleles having repeat lengths of 35 or less. The one false-positive error was due to the previously described potential sample mix-up in 2012B. Among international laboratories, two false-positive errors occurred, yielding a correct response rate of 99.6% (455/457; 95% CI: 98.4–99.9%). These differences in error rates for clinical interpretation also do not differ between US and international laboratories ($P = 0.25$).

DISCUSSION

Both US and international laboratories performed well in the CAP/ACMG external PT challenges. Analytic sensitivity and specificity (a measure of how accurately laboratories could size the trinucleotide repeat length) was over 99% for the US laboratories and over 98% for international laboratories. Overall, 93% of US laboratories (42/45) and 86% of international laboratories (25/29) had no errors in their clinical interpretations. The clinical interpretations were correct for positive challenges between 98 and 99% of the time. Responses for negative challenges were

correct over 99% of the time, regardless of laboratory location. Several incorrect responses were likely transcription errors or sample mix-ups. For example, several false-negative results were identified when the laboratory did not report either a second allele (e.g., “5, blank”) and/or a clinical interpretation was missing. In these instances, the result was considered incorrect because there was no indication of a positive finding. In addition, there were two errors due to sample mix-up.

Challenging clinical laboratories with external PT challenges helps inform the participant, the organizations providing oversight, the government, and the public at large about laboratory test accuracy. Although CAP/ACMG attempt to provide challenges that simulate as closely as possible the usual laboratory clinical scenario, this is not always possible. For example, laboratories usually issue computer-generated reports for clinicians after internal review. For external PT, these reports are most often hand-typed into an online entry form, increasing the chance for typographic and transcription errors. The sizing of mosaic patterns can differ between laboratories or even within a laboratory owing to differences in technicians. Even though the proficiency testing program instructions ask for the size of the densest band, this may differ from routine clinical practice, where the latter is actually reported. On the other hand, the technicians and laboratory directors know this is a PT sample and this alone may lead to better, or different, performance on PT samples versus clinical samples. One limitation of the survey is the lack of any challenges in the “intermediate” range (between 35 and 50 repeats) because these cases are rare and, to date, no appropriate PT material has been identified. However, samples in the congenital range (>1,000 repeats) are now available from the Centers for Disease Control and Prevention Genetic Testing Reference Materials program.¹⁶

These genetic laboratories have demonstrated high performance of their laboratory-developed tests to quantify the number of CTG repeats responsible for DM1 and to provide reliable clinical interpretations. Error rates in US laboratories approach a few per 1,000 challenges, which is equivalent to the highest analytic performance reported in similar surveys.^{17–20} This probably approaches the limit of transcription errors and sample mix-ups, with the very occasional actual false-positive or false-negative analytic result. This high performance is in spite of many laboratories performing the test infrequently (more than half of the US laboratories test fewer than five samples per month) and all tests being developed for use by each laboratory. These results provide strong evidence of the participating laboratory’s ability to provide reliable genetic testing for the diagnosis of DM1.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

The authors thank staff from the College of American Pathologists who supported the Biochemical Molecular Genetics Committee and especially Tom Long, Department of Biostatistics, for reviewing the analyses and Jaimie Gestes Halley, Technical

Manager, Proficiency Testing, for administrative help. The authors also thank Karen Weck, chair of the committee, and other committee members for their helpful comments.

DISCLOSURE

All authors are, or have been, members of the College of American pathologists/American College of Medical Genetics and Genomics (CAP/ACMG) Biochemical and Molecular Genetics Resource Committee. G.E.P. has no other conflicts of interest to declare. M.H. and C.S.R. direct laboratories providing DM1 clinical testing.

REFERENCES

1. Online Mendelian Inheritance in Man. Johns Hopkins University: Baltimore, MD. MIM 160900: 2/21/14. <http://omim.org/>. Accessed 22 February 2016.
2. Bird TD. Myotonic dystrophy type 1. Updated 16 May 2013. In: Pagon RA, Adam MP, Ardinger HH, *et al*. *GeneReviews*. University of Washington: Seattle, WA, 1993–2014. <http://www.ncbi.nlm.nih.gov/books/NBK1165/>. Accessed 22 February 2016.
3. New nomenclature and DNA testing guidelines for myotonic dystrophy type 1 (DM1). The International Myotonic Dystrophy Consortium (IDMC). *Neurology* 2000;54:1218–1221.
4. Moxley RT, Meola G. The myotonic dystrophies. In: *The Molecular and Genetic Basis of Neurologic and Psychiatric Disease*. Wolters Kluwer: Philadelphia, PA, 2008:532–541.
5. Prior TW; American College of Medical Genetics (ACMG) Laboratory Quality Assurance Committee. Technical standards and guidelines for myotonic dystrophy type 1 testing. *Genet Med* 2009;11:552–555.
6. Kamsteeg EJ, Kress W, Catalli C, *et al*. Best practice guidelines and recommendations on the molecular diagnosis of myotonic dystrophy types 1 and 2. *Eur J Hum Genet* 2012;20:1203–1208.
7. Udd B, Krahe R. The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol* 2012;11:891–905.
8. Savkur RS, Philips AV, Cooper TA. Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* 2001;29:40–47.
9. Mankodi A, Takahashi MP, Jiang H, *et al*. Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* 2002;10:35–44.
10. Kanadia RN, Johnstone KA, Mankodi A, *et al*. A muscleblind knockout model for myotonic dystrophy. *Science* 2003;302:1978–1980.
11. Ranum LP, Day JW. Myotonic dystrophy: RNA pathogenesis comes into focus. *Am J Hum Genet* 2004;74:793–804.
12. Day JW, Ranum LP. RNA pathogenesis of the myotonic dystrophies. *Neuromuscul Disord* 2005;15:5–16.
13. Cooper TA. A reversal of misfortune for myotonic dystrophy? *N Engl J Med* 2006;355:1825–1827.
14. Leroy O, Wang J, Maurage CA, *et al*. Brain-specific change in alternative splicing of Tau exon 6 in myotonic dystrophy type 1. *Biochim Biophys Acta* 2006;1762:460–467.
15. Wheeler TM, Thornton CA. Myotonic dystrophy: RNA-mediated muscle disease. *Curr Opin Neurol* 2007;20:572–576.
16. Kalman L, Tarleton J, Hitch M, *et al*. Development of a genomic DNA reference material panel for myotonic dystrophy type 1 (DM1) genetic testing. *J Mol Diagn* 2013;15:518–525.
17. Lyon E, Schrijver I, Weck KE, Ferreira-Gonzalez A, Richards CS, Palomaki GE; CAP/ACMG Biochemical and Molecular Genetics Committee. Molecular genetic testing for cystic fibrosis: laboratory performance on the College of American Pathologists external proficiency surveys. *Genet Med* 2015;17:219–225.
18. Feldman GL, Schrijver I, Lyon E, Palomaki GE; CAP/ACMG Biochemical and Molecular Genetics Resource Committee. Results of the College of American Pathology/American College of Medical Genetics and Genomics external proficiency testing from 2006 to 2013 for three conditions prevalent in the Ashkenazi Jewish population. *Genet Med* 2014;16:695–702.
19. Weck KE, Zehnbauser B, Datto M, Schrijver I; CAP/ACMG Biochemical and Molecular Genetics Resource Committee. Molecular genetic testing for fragile X syndrome: laboratory performance on the College of American Pathologists proficiency surveys (2001–2009). *Genet Med* 2012;14:306–312.
20. Tafe LJ, Datto MB, Palomaki GE, Lacbawan FL; CAP/ACMG Biochemical and Molecular Genetics Resource Committee. Molecular testing for the BRCA1 and BRCA2 Ashkenazi Jewish founder mutations: a report on the College of American Pathologists proficiency testing surveys. *Genet Med* 2015;17:58–62.