Highly accurate molecular genetic testing for *HFE* hereditary hemochromatosis: results from 10 years of blinded proficiency surveys by the College of American Pathologists

Richard D. Press, MD, PhD[1,2,3], Garrett Eickelberg, BAS[1,3], Thomas J. McDonald, MD[4], Jaimie Halley, BS[5], Thomas Long, MPH[6], Laura J. Tafe, MD[7,8] and Karen E. Weck, MD[9,10]; for the CAP/ACMG Biochemical and Molecular Genetics Resource Committee

Purpose: The College of American Pathologists offers blinded proficiency testing (PT) for laboratories performing *HFE* genetic tests for hereditary hemochromatosis (common C282Y and H63D variants). This study used 10 years of PT data to determine laboratory performance for *HFE* analytical genotyping and clinical interpretation.

Methods: Laboratories were graded for accuracy of genotype determination (six possible C282Y/H63D genotypes) and clinical interpretation regarding whether the genotype was likely to have contributed to iron overload in a hypothetical patient.

Results: The analytical genotyping error rate was low (0.73%) in 7,663 results (from 257 unique laboratories). Genotyping errors were significantly higher in C282Y heterozygous, H63D homozygous, and C282Y homozygous samples, in non-American laboratories, and in laboratories with lower testing volume. Analytical sensitivity and

specificity were >98.5 and >99.5%. The interpretive error rate (4.3%) was higher than the genotyping error rate, with two problematic genotypes (C282Y heterozygous and H63D homozygous) accounting for 77% of total interpretive errors. There was a time-dependent improvement in the interpretation of the clinical significance of *HFE* genotypes.

Conclusions: *HFE* molecular genetic testing, performed by non–US Food and Drug Administration–approved laboratory-developed tests, demonstrated excellent accuracy, sensitivity, and specificity. Clinical interpretations were more heterogeneous, probably owing to the low clinical penetrance of some common *HFE* genotypes.

Genet Med advance online publication 28 April 2016

Key Words: diagnostic errors; genetic testing; hemochromatosis; HFE; laboratory proficiency testing

INTRODUCTION

Genetic sequence variation in the *HFE* (high Fe) gene is associated with the most common form of hereditary hemochromatosis (HH), type 1 hemochromatosis. HH can result in severe multiorgan dysfunction due to iron deposition, including cirrhosis, heart disease, and diabetes. *HFE* encodes an MHC class 1–like membrane protein that interacts with β2-microglobulin and transferrin receptor 1 (TfR1), and it is involved in modulating iron absorption by altering the affinity of TfR1 for transferrin-bound iron.[1] Homozygosity for the *HFE* C282Y (c.845G>A) variant is seen in ~80% of individuals with HH, although there is incomplete penetrance.[2–6] The *HFE* C282Y variant disrupts disulfide bridges in the extracellular domains, thus preventing association with β2-microblobulin and TfR1. *HFE* H63D (c.187C>G) has also been associated with iron overload, primarily in C282Y/H63D compound heterozygotes, although the penetrance is very low.[4] Both of these variants are common

in Caucasians, with allele frequencies of ~5–7% (C282Y) and ~13–15% (H63D).[3,7]

Laboratory testing for the *HFE* genotypes associated with HH became one of the most commonly performed molecular genetic tests soon after the identification of *HFE*-associated hemochromatosis. However, the clinical utility of *HFE* genotyping has been somewhat controversial, due primarily to the low penetrance and high mutant allele frequencies of *HFE* sequence variants associated with hemochromatosis. *HFE* genotype testing is often targeted to individuals with suspected iron overload, such as those with increased transferrin saturation and/ or serum ferritin, and it is not recommended as a population screen.[8–11] *HFE* genotype testing can also be useful in individuals with a family history of hemochromatosis. Biochemical iron testing is recommended for all individuals being evaluated for hemochromatosis. Phlebotomy and/or iron chelation therapy are effective in reducing iron overload in at-risk individuals

[1]Department of Pathology, Oregon Health and Science University, Portland, Oregon, USA; [2]Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon, USA; [3]Knight Cancer Institute, Oregon Health and Science University, Portland, Oregon, USA; [4]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts, USA; [5]College of American Pathologists, Northfield, Illinois, USA; [6]Department of Biostatistics, College of American Pathologists, Northfield, Illinois, USA; [7]Dartmouth–Hitchcock Medical Center, Lebanon, New Hampshire, USA; [8]The Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA; [9]Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA; [10]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. Correspondence: Richard D. Press (pressr@ohsu.edu)

with documented iron overload and can improve liver and cardiac function.

The College of American Pathologists (CAP) offers external proficiency testing (PT) for clinical laboratories performing molecular genetic tests, including *HFE* genotyping for the C282Y and H63D variants. The CAP/American College of Medical Genetics and Genomics Biochemical and Molecular Genetics resource committee oversees PT for *HFE* genotyping. Blinded DNA samples selected for carrying different *HFE* genotypes are distributed biannually to participating *HFE*-testing laboratories that then perform *HFE* genotyping using their routine laboratory-developed testing methods. Results are graded by CAP for the accuracy of both the genotype determinations and the clinical interpretations of these genotypes. This report summarizes the cumulative results of laboratory performance for *HFE* analytical genotyping and clinical interpretation by laboratories participating in the CAP PT surveys from 2004 through 2013, covering more than 7,000 reported results from more than 250 different laboratories.

MATERIALS AND METHODS

Samples tested

CAP PT for *HFE* analytical genotyping and clinical interpretation was analyzed over a 10-year period (2004 through 2013), with three blinded samples per survey mailing, two survey mailings per year, and 57 total distributed samples. This PT survey targeted the common C282Y and H63D variants. The CAP-distributed samples were genomic DNA extracted from human cell lines with known *HFE* genotypes (from Coriell). CAP-designed result forms were completed by participating laboratories with analytical and interpretive responses, and returned to CAP for grading. The result forms also contained background questions regarding test method, *HFE* testing volume, and turnaround time for each laboratory. Responses were collated by the CAP and provided blindly to the Biochemical and Molecular Genetics Committee for analysis.

Analysis of *HFE* analytical genotyping

There were six potential genotypes that could be reported by participating laboratories: homozygous wild-type, heterozygous C282Y, heterozygous H63D, compound heterozygous C282Y/H63D, homozygous C282Y, and homozygous H63D. If the reported genotype did not agree with the known *HFE* genotype of the distributed cell line, then it was considered incorrect.

Analysis of clinical interpretation

The laboratories that provided a correct analytical *HFE* genotype were also graded on their clinical interpretation for that genotype result, given the clinical scenario that the sample was from a patient with documented biochemical iron overload. For survey years 2005 to 2013, the three possible interpretive responses (from which laboratories were asked to choose a single best response) were "This genotype is likely contributing to iron overload," "There is no evidence that this genotype contributes significantly to iron overload," or "Inconclusive." For survey year 2004 (only), the third interpretive choice was "This genotype may or may not contribute to iron overload," which was subsequently changed to "Inconclusive" beginning in 2005. For the purposes of this study, the consensus acceptable interpretations for each genotype (by which laboratories were graded) are detailed in **Table 1**. Any reported clinical interpretations other than those listed in **Table 1** were considered unacceptable responses.

Data analysis

To remove potential grading bias, the genotypes and interpretations reported at the time of the initial survey were re-graded for every sample. The analytical and interpretive grades were also compared across various other practices, laboratories, or demographic variables. Chi-squared tests of association were used to compare the grades across the various levels of the comparison variables. In addition to the overall chi-squared test, each level of the comparison variable was simultaneously tested (chi-squared) against the other remaining levels combined. The significance level was corrected using the Bonferroni correction factor. For ordered categories of selected comparison variables (i.e., *HFE* testing volume, survey year), a test of trends was conducted using the Cochran-Armitage test. Confidence limits were calculated using the exact quadratic formula. Statistical analyses were performed with SAS software.

RESULTS

Between 2004 and 2013, a cumulative total of 257 unique laboratories participated in the CAP *HFE* proficiency survey. On average, 138 (SE: 1.8) laboratories subscribed to each biannual survey, and each laboratory received an average of 30 cumulative samples (SE: 1.3). The total number of distributed samples within each of the six genotype groups is detailed in **Table 2**. The genotyping methods used by participating laboratories

Table 1 Consensus clinical interpretations (for grading)

Genotype	Best acceptable interpretation	Other acceptable interpretation
Homozygous C282Y	Genotype likely contributes to iron overload	
Compound heterozygous (C282Y/H63D)	Genotype likely contributes to iron overload	Inconclusive[a]
Homozygous H63D	Genotype likely contributes to iron overload	Inconclusive[a]
Heterozygous C282Y	No evidence genotype contributes significantly to iron overload	—
Heterozygous H63D	No evidence genotype contributes significantly to iron overload	—
Wild-type	No evidence genotype contributes significantly to iron overload	—

[a]The interpretation "Genotype may or may not contribute to iron overload" was used in a single 2004 survey and was considered equivalent to the subsequent "inconclusive" interpretive category.

Table 2 Analytical errors across all genotypes

	Homozygous C282Y	Compound heterozygous	Homozygous H63D	Heterozygous C282Y	Heterozygous H63D	Wild-type	Total
Correct results	797	805	273	526	525	4,681	7,607
Incorrect results	11	5	7	9	5	19	56
Total samples	808	810	280	535	530	4,700	7,663
Error percentage	1.4	0.62	2.5	1.7	1.0	0.40	0.73
P (vs. other combined groups)	0.026	NS	0.0004	0.007	NS	<0.0001	—
Number of times included in survey	6	6	2	4	4	35	57

Table 3 Analytical sensitivity and specificity of *HFE* genotyping

Assumed detection target	Analytical sensitivity (95% CI)	Analytical specificity (95% CI)
Any non-wild-type[a]	99.39% (99.02–99.63)	99.60% (99.36–99.75)
Homozygous C282Y[b]	98.64% (97.50–99.28)	99.93% (99.82–99.97)

[a]Correctly genotyped *HFE* wild-type samples were defined as "true negative" (specificity = 4,681/4,700), and samples correctly identified as having any *HFE* mutation were defined as "true positive" (sensitivity = 2,945/2,963, including 19 samples with the wrong mutant genotype). [b]Correctly genotyped C282Y homozygous samples were defined as "true positive" (sensitivity = 797/808), and samples correctly identified as having any genotype other than homozygous C282Y were defined as "true negative" (specificity = 6,850/6,855, including 40 samples with the wrong non-C282Y/C282Y genotype).

were quite diverse, with the most popular methods being differential restriction enzyme fragment lengths (51%), melting curve-based methods (15%), probe-specific real-time PCR (TaqMan) methods (8%), direct sequencing (7%), and allele-specific PCR (7%).

Analytical genotype performance

Of the 7,663 total samples graded for the *HFE* analytical genotype across the 10-year survey period, there were 56 analytical errors across all laboratories and all genotypes (error rate, 0.73%; 95% confidence interval (CI): 0.54–0.99%) (**Table 2**). Of these 7,663 blinded PT samples, 99.3% were determined to have the correct *HFE* analytical test result (95% CI: 99.01–99.46%). By examining the temporal pattern of correct and incorrect analytical genotype results across each laboratory and survey, we could potentially attribute approximately half (*n* = 27) of the 56 total errors to a possible preanalytical sample switch with another sample from the same survey rather than to a true analytical error. There were 12 possible two-way sample switches and one possible three-way sample switch.

There was a significantly higher analytical error rate for C282Y heterozygous (1.7%) and H63D homozygous (2.5%) samples, and a trend toward a higher error rate for C282Y homozygous samples (1.4%) as compared with the aggregate of the other five genotype groups. There was a significantly lower analytical error rate for wild-type samples (0.4%) as compared with the aggregate of the other genotypes.

Of the 257 unique laboratory participants, 219 (85%) had no analytical errors over the entire duration of the laboratory's participation in this survey (average 28 consecutive concordant samples per laboratory). Of the 38 laboratories (15%) that

did make a genotype error at some time during the 10-year analysis, 14 laboratories (5.4%) made more than one genotype error (accounting for 32 cumulative errors) and 24 laboratories (9.3%) made a single error. A small fraction of laboratories (*n* = 14; 5.4%) was therefore responsible for a majority (32/56; 57%) of the total analytical genotype errors.

The analytical error rate was significantly higher among the 1,398 results from laboratories not located within North America (23 errors; 1.6%) than among the 6,229 results from North American (US and Canadian) laboratories (33 errors; 0.53%) (*P* < 0.0001). Additionally, the United States–based laboratories (*n* = 5,197 samples) had a lower error rate (0.54%) compared with the non-US laboratories (*n* = 2,430 samples; 1.15%) (*P* = 0.003). There was no significant difference in the analytical error rate among laboratories using any particular genotyping method. In particular, the 51% of laboratories using differential restriction enzyme fragment lengths as their genotyping method (by far the most popular method) had a 0.73% analytical genotyping error rate compared with a cumulative 0.75% error rate for laboratories using any other nonrestriction enzyme-based methodology (*P* ≥ 0.9). The method-specific error rates for these nonrestriction enzyme-based techniques were 0.54% for melting curve–based methods, 0.66% for probe-specific real-time PCR (TaqMan) methods, 0.59% for direct sequencing, and 1.5% for allele-specific PCR. None of these method-specific error rates was statistically significantly different from the others (*P* ≥ 0.4).

As part of the CAP survey, laboratories were asked to self-report their *HFE* testing volume. To assess the relationship between analytical error rate and number of monthly samples tested by a laboratory, we defined four unbiased, approximately equally sized, quartile groups of test volume (1–6, 7–15, 16–40, and >40 tests per month). Using a Cochran-Armitage trend test in these four quartile-defined groups, the analytical error rate progressively decreased as the test volume progressively increased—with error rates of 1.1, 0.80, 0.67, and 0.49%, respectively, for the lowest to highest volume quartiles (*P* = 0.044).

Sensitivity and specificity of analytical genotyping
HFE genotyping analytical sensitivity and specificity were calculated based on the 7,663 total genotype results obtained over the cumulative 10-year period (**Table 3**). For these sensitivity and specificity calculations, two different assumptions were made regarding the clinically relevant "true positive" genotype intended to be detected by this test: (i) any non-wild-type

sample or (ii) only a C282Y homozygous sample. The analytical sensitivity and specificity were >98.5 and >99.5%, respectively, for detecting any abnormal *HFE* genotype and for detecting only the highest-risk homozygous C282Y genotype.

Clinical interpretation performance

For each blinded survey sample, laboratories were also asked to interpret the clinical relevance of their genotype result for a hypothetical patient with biochemical iron overload. Overall interpretive error rates were calculated over the 10-year period based on the acceptable interpretive response for each genotype as defined in **Table 1**. Of the 7,562 samples with an evaluable clinical interpretation, 7,240 samples (95.7%) had an acceptable interpretive response and 322 samples (4.3%) had an unacceptable interpretive response (**Table 4**). This 4.3% interpretive error rate (95% CI: 3.9–4.8%) was significantly greater than the observed analytical genotyping error rate of 0.73% rate (95% CI: 0.54–0.99%) ($P < 0.001$). In contrast to the 85% (219/257) of laboratories that never had an analytical genotyping error over their entire time period of survey participation, only 34% of laboratories ($n = 87$) gave a consistently correct interpretive response for every analyzed sample.

A genotype-specific breakdown of the interpretive error rates (**Table 4** and **Figure 1**) shows that >94% of interpretive responses were acceptable for the homozygous wild-type, heterozygous H63D, compound heterozygous C282Y/H63D, and homozygous C282Y genotypes. The C282Y heterozygous and H63D homozygous samples, by contrast, yielded an interpretive error rate much higher (24 and 46%, respectively) than that for the other genotype groups ($P < 0.0001$). These two error-prone genotype groups cumulatively accounted for 77% of the total interpretive errors (249/322). Conversely, homozygous C282Y (0.13%) and homozygous wild-type (0.19%) samples had a statistically lower interpretive error rate than the other genotype groups ($P < 0.0001$).

The distribution of each of the three possible interpretive responses given for each of the six genotype groups is shown in **Figure 1** ($n = 7,562$). In three genotype groups (homozygous wild-type, heterozygous H63D, and homozygous C282Y), there was a high level of consensus, with a single best interpretive response given by 99.8, 94.5, and 99.9% of laboratories, respectively. For the compound heterozygous C282Y/H63D genotype, although only 4.3% of responses were considered unacceptable ("no evidence that the genotype contributes significantly to iron overload"), there was substantial heterogeneity in the two acceptable responses: 77% of interpretive responses were that

the C282Y/H63D genotype "likely contributes to iron overload" (the best response), and 19% of interpretive responses were that the clinical significance of this genotype is "inconclusive." For C282Y heterozygous samples, 76% of responses were the best acceptable interpretation that this genotype does not contribute significantly to iron overload. However, a cumulative 24% of unacceptable responses were that the clinical relevance of this genotype is either "inconclusive" (13% of responses) or that it "likely contributes" to iron overload (11% of responses). The H63D/H63D homozygous genotype had the greatest heterogeneity in clinical interpretation, with only 29% of respondents interpreting that in an individual with documented iron overload "This genotype likely contributes to iron overload," which was considered the best acceptable interpretive response. Another 25% of responses were that the clinical correlative evidence for the H63D homozygous genotype is "inconclusive" (also considered an acceptable response). However, almost half (46%) of the total responses were that there is no evidence that the H63D homozygous genotype contributes to iron overload, which was considered an unacceptable interpretation.

In an attempt to determine whether laboratories with unacceptable interpretive responses subsequently improved their interpretive accuracy after participation in the education-focused CAP survey, we examined the distribution of interpretive responses over time for the three error-prone genotypes (**Figure 2**). In the four sequential surveys with a C282Y heterozygous sample, although the cumulative error rate was high (24%), the percentage of laboratories responding with the best acceptable interpretation that there is "no evidence that the genotype contributes to iron overload" progressively improved from 65% in 2005 to 82% in 2011 (P for trend < 0.001) (**Figure 2a**). In the two surveys with an H63D homozygous sample, the percentage of laboratories responding with the best acceptable interpretation that this genotype is "likely contributing to iron overload" increased from 23 to 34% from 2009 to 2013 ($P = 0.04$) (**Figure 2b**). For the compound heterozygous C282Y/H63D genotype, the percentage of laboratories correctly interpreting this genotype as "likely contributing to iron overload" increased from 34% in 2004 to 96% in 2013 (P for trend < 0.001) (**Figure 2c**).

DISCUSSION

Since the 1996 discovery[12] that *HFE* C282Y homozygosity accounts for the overwhelming majority of HH cases in Caucasians, the CAP *HFE* PT survey has had one of the largest enrollment rates, with more than 250 unique laboratory

Table 4 Interpretive responses across all genotypes

	Homozygous C282Y	Compound heterozygous	Homozygous H63D	Heterozygous C282Y	Heterozygous H63D	Wild-type	Total
Acceptable response	790	765	147	400	494	4,644	7,240
Unacceptable response	1	34	125	124	29	9	322
Total samples	791	799	272	524	523	4,653	7,562
Error percentage	0.13	4.3	46	24	5.5	0.19	4.3
P value (vs. other groups)	<0.0001	NS	<0.0001	<0.0001	NS	<0.0001	—
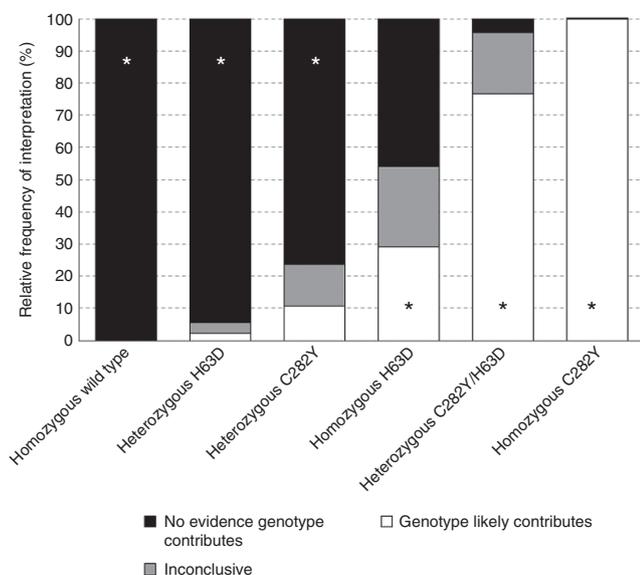
Figure 1 Relative distribution of interpretive responses (for genotype contribution to iron overload) across all genotypes and samples (total n = 7,562). *Interpretations graded as the best acceptable response. For the H63D homozygous and compound heterozygous (C282Y/H63D) genotypes, "inconclusive" was also considered an acceptable response.

participants over the 2004 to 2013 time period. A prior publication summarized the accuracy for detecting only the *HFE* C282Y homozygous genotype in the 1998–2002 CAP PT surveys.[13] We have now expanded the analysis of laboratory performance over a longer and more recent 10-year period, and we have assessed the analytical genotyping accuracy of both common *HFE* variants (including H63D) in addition to the clinical interpretation of these genotype results.

The overall error rate for genotype determinations was low (0.73%), with an analytical sensitivity of >98.5% and specificity of >99.5%. Furthermore, 85% of participating laboratories had a perfect genotyping record (no errors) over this entire decade of testing. This represents excellent analytical sensitivity and specificity, with few false-positive or false-negative results by laboratories performing *HFE* genotyping over these 10 years. Although the overall genotyping accuracy was high, with an error rate of <3% for all *HFE* genotypes, the C282Y heterozygous and H63D homozygous genotypes demonstrated significantly higher analytical errors, and the homozygous wild-type genotype demonstrated significantly lower errors compared with all other combined genotype groups. Our updated analysis of analytical sensitivity and specificity for detecting the highest-risk *HFE* C282Y homozygous genotype yielded similar results as the previous analysis of CAP PT data from 1998 to 2002, which found a sensitivity of 98.4% (95% CI: 95.9–99.5) and a specificity of 99.8% (95% CI: 99.4–99.9) over this initial non-overlapping 5-year period.[13] These data indicate that the analytical accuracy of *HFE* testing laboratories has remained very high over the past 15 years.

Nonetheless, any false-positive results will result in a positive predictive value of less than 100%, which would have

significant implications for population screening.[13] Although population screening for *HFE* is generally not recommended, the CAP specificity data indicate that screening could result in 3–18 (95% CI) false-positive results for every 10,000 individuals, assuming that laboratory performance for PT is an accurate indicator of overall testing accuracy. Thus, in an asymptomatic individual with an isolated *HFE* C282Y homozygous result, confirmation of risk of iron overload should be confirmed by secondary testing and/or by determining biochemical indices of iron overload. In a patient who underwent *HFE* testing with confirmed biochemical iron overload, a false-negative result could lead to additional unnecessary testing and the ultimate failure to attribute this iron overload to an inherited condition that can be successfully treated with phlebotomy. Progressive clinical symptoms and end-organ damage could ensue. A false-positive *HFE* testing result could, by contrast, result in unnecessary phlebotomy. Although this is not a particularly toxic treatment, it is inconvenient and, for nonhemochromatosis iron overload, usually ineffective. The familial consequences of inaccurate *HFE* genetic testing could also be substantial and might include delayed (or absent) case discovery in other family members. This could delay or prevent effective therapeutic intervention before the onset of irreversible end-organ damage.

The *HFE* genotyping error rates determined from CAP PT data may be an overestimate of the overall error rate of clinical *HFE* genotyping performed in US diagnostic laboratories. In general, North American (US and Canadian) laboratories performed better than non–North American laboratories with respect to genotype accuracy, and higher-volume laboratories performed better than those with a lower volume of testing. However, in all of these laboratory subcategories, the analytical error rate was less than 2%. In addition, preanalytic sample switching could have possibly accounted for a significant proportion of errors (~50%) observed in these PT surveys, based on a comparison of the expected genotype for each of the three samples concurrently tested in a given survey. An analogous observation has been documented for other CAP PT surveys,[14,15] suggesting that analytical genotyping accuracy for real-world clinical samples may be higher than that measured by the PT surveys. Although there is no objective way of determining whether preanalytic errors are higher for the PT survey samples than for clinical samples, it is worth noting that due to the nature of the PT survey samples, they are typically not barcoded or labeled in the same manner that clinical samples are and therefore must be processed and handled differently (manually), which may lead to an increased rate of preanalytic labeling errors. Although laboratories are aware that these non-routine samples are CAP proficiency samples, they are completely blinded regarding the genotype results.

In addition to analyzing analytical accuracy, we also analyzed the accuracy of the clinical interpretation of the genotype results using a clinical scenario of a patient with documented biochemical iron overload, which is the recommended clinical indication for *HFE* genetic testing.[9–11] Overall interpretive errors (4.3%) were significantly higher than analytical genotype
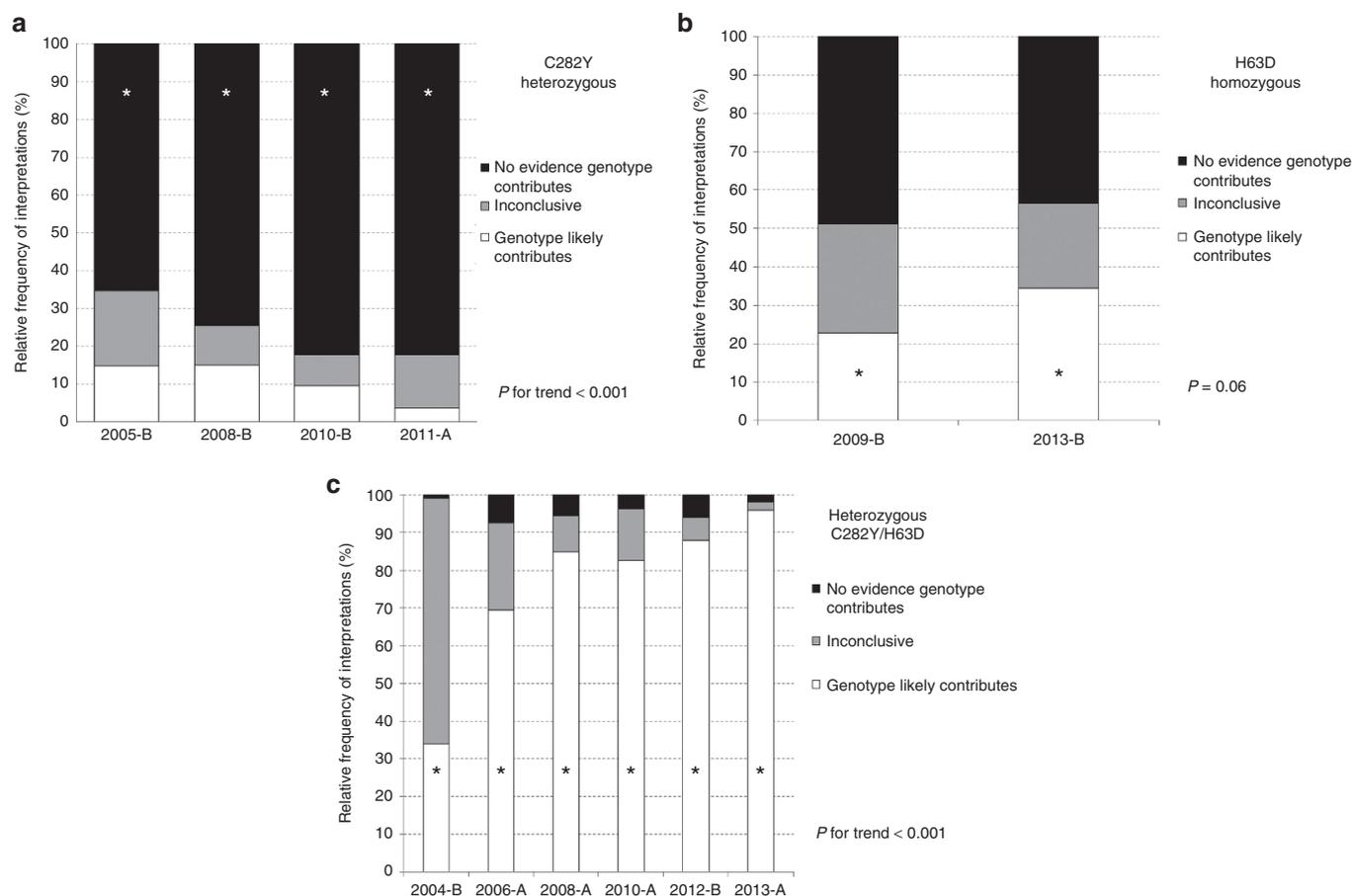
Figure 2 Relative distributions of interpretive responses (for genotype contribution to iron overload) for the (a) C282Y heterozygous, (b) H63D homozygous, and (c) compound heterozygous C282Y/H63D genotype in every individual survey sample with these genotypes. Survey years ending in "-A" or "-B" indicate the first or second biannual survey, respectively. *The best acceptable response. For the H63D homozygous and compound heterozygous genotypes, "inconclusive" was also considered an acceptable response.

errors. Of note, 66% of laboratories made at least one interpretive error that was in variance from the intended interpretations. The interpretation of two genotypes stood out as particularly problematic: C282Y heterozygous and H63D homozygous. There was also variability, although to a lesser degree, in the clinical interpretation of iron-overloaded subjects with the compound heterozygous C282Y/H63D genotype. The lack of consensus for these three genotypes is possibly due to conflicting literature regarding the relative risk and penetrance of these genotypes, and/or to failure to appreciate the difference in interpreting results in an individual with documented biochemical iron overload rather than screening healthy individuals.

Many previous studies—several in large screening populations—have addressed the question of the complex genotype–phenotype correlations between the two common *HFE* variants (C282Y and H3D), the clinical syndrome of symptomatic hemochromatosis, and/or its asymptomatic iron overload predecessor state. All of these studies have universally confirmed that the homozygous C282Y genotype is both the overwhelmingly predominant genotype in patients with clinically symptomatic hemochromatosis and the genotype predisposing to

the highest risk of present and future progressive biochemical iron overload in subjects with no current symptomatic disease.[2–4,6] The *HFE* H63D variant, in both the homozygous and compound heterozygous C282Y/H63D state, is associated with a much lower penetrance than homozygous C282Y with respect to the frequency and magnitude of elevated iron indices (i.e., ferritin and transferrin saturation) as well as to the risk of developing clinical symptoms and hemochromatosis-related end-organ disease.[4–6,16,17] Additional modifying risk factors likely play a major role in those rare H63D individuals who do develop iron overload.[8,18] Thus, the observed heterogeneity in interpretations for the H63D homozygous and compound heterozygous C282Y/H63D genotypes may be due to the failure to recognize that, even though the penetrance for these genotypes is low, in an individual with documented iron overload, it is likely that these genotypes are contributing to the phenotype. Nevertheless, given the somewhat conflicting evidence regarding the disease risk and clinical penetrance of the H63D variant, the "inconclusive" interpretive response was considered acceptable for these two low-penetrance genotypes for all analyses in this article.

HH is an autosomal-recessive disorder. Therefore, heterozygosity for C282Y is not associated with a significant risk of developing clinical hemochromatosis. Nevertheless, 4–15% of laboratories responded that *HFE* C282Y heterozygosity would "likely contribute" to biochemical iron overload. Of note, an additional 8–20% of laboratories responded that the C282Y heterozygous genotype was "inconclusive" for contributing to iron overload. This observed lack of consensus in interpretive responses may be due to some published population screening studies showing an increased trend in ferritin and transferrin saturation parameters in C282Y heterozygotes compared with wild-type individuals.[2,19] However, the range of iron indices in these studies overlapped with those in wild-type individuals, and the overwhelming evidence indicates that C282Y heterozygotes are not at increased risk of developing clinical hemochromatosis.[3,5,16,20]

Despite the observed lack of consensus, interpretive responses for the problematic genotypes improved over the 10-year time window of evaluation. This performance improvement may be due to prolonged participation in the PT survey (perhaps as a result of education via the PT survey analysis discussion that is distributed to participating laboratories) or to additional clarifying publications in the literature over the time period analyzed. It should also be noted that it is difficult to accurately assess the clinical interpretation of a laboratory result in a simple multiple-choice survey format, and CAP does not solicit laboratory reports for a more complex analysis of reporting language. Patient-specific laboratory reports are likely to use clinical interpretation language that is more nuanced and descriptive than the three simplified discrete choices provided by the CAP survey. A recent publication investigating whether test results for *HFE* C282Y heterozygosity were associated with increased patient anxiety or harm indicated that these results were being appropriately interpreted as associated with no increased risk of disease.[21] However, it is possible that the observed PT survey heterogeneity in clinical interpretations of *HFE* genotyping may also be reflected in clinical diagnostic laboratory test reports that go to clinicians, and some misinterpretation of results may then ensue. There may be a need for further consensus and standardization in this area.

Previous publications have analyzed CAP/ACMG PT data for a variety of molecular genetic tests, including microsatellite instability, Tay-Sachs disease, Canavan disease, familial dysautonomia, *BRCA1/2*, pharmacogenetic analytes, Huntington disease, cystic fibrosis, fragile X syndrome, and other rare inherited disorders.[14,15,22–27] Other publications have examined cystic fibrosis molecular testing in both the United States and Italy.[28,29] All of these publications indicate that molecular genetic tests have excellent analytic performance, with an overall sensitivity greater than 95% and specificity greater than 99%. In addition, these published analytical error rates for molecular genetic tests are quite comparable to those found for other more automated diagnostic analytes in laboratory medicine, with well-documented analytical error rates ranging from 0.01 to 5%.[30–36] Prior publications have also documented findings analogous to this study concerning a higher variability in interpretive (compared with analytical) performance, with interpretive accuracy rates ranging from 92.5% for *BRCA1/2* (ref. 25) to 99.7% for familial dysautonomia.[23] Our findings for *HFE* are consistent with these results, with respect to both analytic and interpretive performance. Virtually all of the aforementioned genetic testing is performed using laboratory-developed tests or laboratory-developed procedures, and there are no commercially available US Food and Drug Administration–cleared tests for *HFE* genotyping. Laboratory-developed molecular methods for the detection of genetic diseases thus yield highly accurate analytical results.

ACKNOWLEDGMENT

DISCLOSURE

REFERENCES

1. Vujić M. Molecular basis of HFE-hemochromatosis. *Front Pharmacol* 2014;5:42.
2. Allen KJ, Gurrin LC, Constantine CC, et al. Iron-overload-related disease in HFE hereditary hemochromatosis. *N Engl J Med* 2008;358:221–230.
3. Adams PC, Reboussin DM, Barton JC, et al.; Hemochromatosis and Iron Overload Screening (HEIRS) Study Research Investigators. Hemochromatosis and iron-overload screening in a racially diverse population. *N Engl J Med* 2005;352:1769–1778.
4. Gallego CJ, Burt A, Sundaresan AS, et al. Penetrance of hemochromatosis in HFE genotypes resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE network. *Am J Hum Genet* 2015;97:512–520.
5. Neghina AM, Anghel A. Hemochromatosis genotypes and risk of iron overload–a meta-analysis. *Ann Epidemiol* 2011;21:1–14.
6. Njajou OT, Houwing-Duistermaat JJ, Osborne RH, et al. A population-based study of the effect of the HFE C282Y and H63D mutations on iron metabolism. *Eur J Hum Genet* 2003;11:225–231.
7. Exome Aggregation Consortium (ExAC). ExAC browser. http://exac.broadinstitute.org; http://exac.broadinstitute.org/gene/ENSG00000010704 (for HFE allele frequencies). Accessed 14 September 2015.
8. Hanson EH, Imperatore G, Burke W. HFE gene and hereditary hemochromatosis: a HuGE review. Human Genome Epidemiology. *Am J Epidemiol* 2001;154:193–206.
9. King C, Barton DE. Best practice guidelines for the molecular genetic diagnosis of type 1 (HFE-related) hereditary haemochromatosis. *BMC Med Genet* 2006;7:81.
10. Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS; American Association for the Study of Liver Diseases. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. *Hepatology* 2011;54:328–343.
11. Porto G, Brissot P, Swinkels DW, et al. EMQN best practice guidelines for the molecular genetic diagnosis of hereditary hemochromatosis (HH). *Eur J Hum Genet* 2016;24:479–495.
12. Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat Genet* 1996;13:399–408.
13. Palomaki GE, Haddow JE, Bradley LA, Richards CS, Stenzel TT, Grody WW. Estimated analytic validity of HFE C282Y mutation testing in population screening: the potential value of confirmatory testing. *Genet Med* 2003;5:440–443.
14. Weck KE, Zehnbauer 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.
15. Lyon E, Schrijver I, Weck KE, Ferreira-Gonzalez A, Richards CS, Palomaki GE; CAP/ACMG Biochemical and Molecular Genetics Committee. Molecular

ORIGINAL RESEARCH ARTICLE

genetic testing for cystic fibrosis: laboratory performance on the College of American Pathologists external proficiency surveys. *Genet Med* 2015;17: 219–225.

16. Zaloumis SG, Allen KJ, Bertalli NA, et al.; HealthIron Study Investigators. Natural history of HFE simple heterozygosity for C282Y and H63D: a prospective 12-year study. *J Gastroenterol Hepatol* 2015;30:719–725.

17. Kelley M, Joshi N, Xie Y, Borgaonkar M. Iron overload is rare in patients homozygous for the H63D mutation. *Can J Gastroenterol Hepatol* 2014;28:198–202.

18. Burke W, Imperatore G, McDonnell SM, Baron RC, Khoury MJ. Contribution of different HFE genotypes to iron overload disease: a pooled analysis. *Genet Med* 2000;2:271–277.

19. Bulaj ZJ, Griffen LM, Jorde LB, Edwards CQ, Kushner JP. Clinical and biochemical abnormalities in people heterozygous for hemochromatosis. *N Engl J Med* 1996;335:1799–1805.

20. Seckington R, Powell L. HFE-associated hereditary hemochromatosis. In: Pagon RA, Adam MP, Ardinger HH, et al. (eds). *GeneReviews*. Seattle, WA, 1993–2015 (updated 17 September 2015).

21. Delatycki MB, Wolthuizen M, Aitken MA, Hickerton C, Metcalfe SA, Allen KJ. To tell or not to tell—what to do about p.C282Y heterozygotes identified by HFE screening. *Clin Genet* 2013;84:286–289.

22. Boyle TA, Bridge JA, Sabatini LM, et al.; College of American Pathologists Molecular Oncology Committee. Summary of microsatellite instability test results from laboratories participating in proficiency surveys: proficiency survey results from 2005 to 2012. *Arch Pathol Lab Med* 2014;138:363–370.

23. 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.

24. Richards CS, Palomaki GE, Lacbawan FL, Lyon E, Feldman GL; CAP/ACMG Biochemical and Molecular Genetics Resource Committee. Three-year experience of a CAP/ACMG methods-based external proficiency testing program for laboratories offering DNA sequencing for rare inherited disorders. *Genet Med* 2014;16:25–32.

25. 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.

26. Palomaki GE, Richards CS. Assessing the analytic validity of molecular testing for Huntington disease using data from an external proficiency testing survey. *Genet Med* 2012;14:69–75.

27. Wu AH. Genotype and phenotype concordance for pharmacogenetic tests through proficiency survey testing. *Arch Pathol Lab Med* 2013;137: 1232–1236.

28. Earley MC, Laxova A, Farrell PM, et al. Implementation of the first worldwide quality assurance program for cystic fibrosis multiple mutation detection in population-based screening. *Clin Chim Acta* 2011;412:1376–1381.

29. Salvatore M, Falbo V, Floridia G, et al. The Italian External Quality Control Programme for cystic fibrosis molecular diagnosis: 4 years of activity. *Clin Chem Lab Med* 2007;45:254–260.

30. Witte DL, VanNess SA, Angstadt DS, Pennell BJ. Errors, mistakes, blunders, outliers, or unacceptable results: how many? *Clin Chem* 1997;43(8 Pt 1): 1352–1356.

31. Carraro P, Plebani M. Errors in a stat laboratory: types and frequencies 10 years later. *Clin Chem* 2007;53:1338–1342.

32. Szecsi PB, Ødum L. Error tracking in a clinical biochemistry laboratory. *Clin Chem Lab Med* 2009;47:1253–1257.

33. Sakyi A, Laing E, Ephraim R, Asibey O, Sadique O. Evaluation of analytical errors in a clinical chemistry laboratory: a 3 year experience. *Ann Med Health Sci Res* 2015;5:8–12.

34. Steindel SJ, Howanitz PJ, Renner SW; College of American Pathologists. Reasons for proficiency testing failures in clinical chemistry and blood gas analysis: a College of American Pathologists Q-Probes study in 665 laboratories. *Arch Pathol Lab Med* 1996;120:1094–1101.

35. Plebani M, Carraro P. Mistakes in a stat laboratory: types and frequency. *Clin Chem* 1997;43(8 Pt 1):1348–1351.

36. Nutting PA, Main DS, Fischer PM, et al. Toward optimal laboratory use. Problems in laboratory testing in primary care. *JAMA* 1996;275:635–639.