

Estimated analytic validity of *HFE* C282Y mutation testing in population screening: The potential value of confirmatory testing

Glenn E. Palomaki, BS¹, James E. Haddow, MD¹, Linda A. Bradley, PhD², C. Sue Richards, PhD³, Timothy T. Stenzel, MD, PhD⁴, and Wayne W. Grody, MD, PhD⁵

Purpose: The purpose of this study was to estimate analytic sensitivity and specificity of *HFE* testing for C282Y homozygosity in the hypothetical setting of population screening for hemochromatosis. **Methods:** We analyzed published results of the Molecular Genetics Survey performed by the American College of Medical Genetics/College of American Pathologists between 1998 and 2002, taking into account its educational nature. **Results:** Analytic sensitivity for C282Y homozygosity is 98.4% (95% CI 95.9%–99.5%). The analytic specificity is 99.8% (99.4%–99.9%). At a frequency of 40 per 10,000 for the homozygous genotype, the analytic positive predictive value is 66%. **Conclusion:** *HFE* testing for C282Y homozygosity is highly reliable. Homozygosity is uncommon in population screening, however, and confirmatory testing should be considered. *Genet Med* 2003;5(6):440–443.

Key Words: hemochromatosis, analytic validity, *HFE*, C282Y, proficiency testing

The term analytic validity refers to a laboratory's documented ability to accurately perform a given measurement and report the result. Documenting analytic validity is a necessary component of a reliable genetic test.¹ Analytic validity includes analytic sensitivity, analytic specificity, analytic predictive power, assay robustness, and appropriate quality assurance measures. The present study addresses measurement of the C282Y mutation of the *HFE* gene and assesses the first three components (analytic sensitivity, specificity, and predictive power). Analytic sensitivity is defined as the proportion of positive test results correctly reported by the laboratory among samples containing a mutation that the laboratory's test is designed to detect. Analytic specificity is defined as the proportion of negative test results correctly reported by the laboratory when no detectable mutation is present. The analytic false-positive rate (1-analytic specificity) is another way to express analytic specificity. The analytic positive predictive power is defined as the proportion of positive test results that are correct.

Estimates of analytic performance can vary depending upon the way in which the test is used. Currently, clinical testing involving *HFE* mutations is performed mainly in symptomatic individuals, or in those with a family member either diagnosed

with hemochromatosis, or with identified *HFE* mutations. The present study looks beyond these current indications for testing and explores the test performance in the theoretical context of population screening. The *HFE* genotype with the highest penetrance for symptomatic iron overload is C282Y homozygosity.^{2,3} It is not yet clear whether the penetrance of even this genotype is sufficiently high to justify its use in population screening.⁴ Other *HFE* genotypes are of much lower penetrance and do not warrant consideration in a screening setting.^{2,3} For that reason, the present study examines the screening implications of only the C282Y genotype, using data collected via external proficiency testing from laboratories presently performing diagnostic testing.

MATERIALS AND METHODS

External proficiency testing for *HFE* mutations is jointly sponsored by the American College of Medical Genetics (ACMG) and the College of American Pathologists (CAP). Purified DNA from established cell lines is distributed once or twice a year to participating laboratories. The semiannual reports available from the Molecular Genetics Resource Committee are the source for the data used in the present analysis.⁵ The strengths and weakness of using this data source versus other data sources have been previously described.⁶ The ACMG/CAP survey includes nearly all clinical laboratories in the U.S. that utilize a wide range of *HFE* testing methodologies. The survey samples have confirmed genotypes and are tested blindly. However, basing analytic performance estimates on the ACMG/CAP program data also has drawbacks. These include the overrepresentation of "difficult" samples due to the educational nature of the program, mixing of screening and diagnostic exercises, the "artificial" nature of sample prepara-

From the ¹Foundation for Blood Research, Scarborough, Maine; ²Centers for Disease Control and Prevention, Atlanta, Georgia; ³Baylor College of Medicine, Houston, Texas; ⁴Duke University Medical Center, Durham, North Carolina; and ⁵UCLA School of Medicine, Los Angeles, California.

Glenn E. Palomaki, Foundation for Blood Research, PO Box 190, 69 US Route 1, Scarborough, ME 04070-0190.

Received: July 9, 2003.

Accepted: September 8, 2003.

DOI: 10.1097/01.GIM.0000096500.66084.85

tion, shipping and handling, and the inclusion of laboratories from outside the U.S., as well as reagent manufacturers or research laboratories.

One additional consideration might be that laboratories perform differently when testing proficiency samples than when routinely testing clinical samples, even though CLIA regulations require proficiency samples to be tested in the same manner as patient samples. The performance might be better because special attention might be given to the sample by the laboratory. Alternatively, the performance might be less good, because the sample could not be processed according to the routine laboratory protocol (e.g., the original sample is extracted DNA rather than blood or buccal scrapings).

Although the majority of participating laboratories tested for both the C282Y and H63D mutation, only results pertaining to C282Y are included in the present analysis (for analyses of other genotypes/mutations, see <http://www.cdc.gov/genomics/activities/FBR/HH/HHAnaVal.htm>). For example, if a compound heterozygous sample (C282Y/H63D) was distributed, it was analyzed as if it were heterozygous only for C282Y (C282Y/wild). The present study focuses on the analytic performance of HFE testing and estimates the analytic positive predictive power (the proportion of reported C282Y homozygotes that are correct). It is reasonable to expect that this rate should approach 100%. The clinical positive predictive power (the proportion of C282Y homozygotes that will develop clinical manifestations of hemochromatosis) will be much lower because of incomplete penetrance. Ninety-five percent confidence intervals are computed using the binomial distribution (True EPISTAT, Richardson, Texas).

RESULTS

From 1998 through 2002, between 67 and 103 laboratories participated in the ACMG/CAP Molecular Genetics Laboratory (MGL) survey.⁵ Participants included both clinical and nonclinical laboratories from the United States and elsewhere and utilized a wide range of analytic techniques. In 2002, for example, 103 participating laboratories reported results for three sample challenges in the spring, and 98 participating laboratories reported results for three additional sample challenges in the fall. Thus, there were 603 sample challenges for which laboratories reported an HFE genotype. Table 1 shows

Table 1.

HFE C282Y mutation testing: A summary of ACMG/CAP molecular genetics survey results for 1998–2002

Test result	Actual genotype ^a			Total
	C282Y/C282Y	C282Y/wild	Wild/wild	
C282Y/C282Y	243	1	3	247
C282Y/wild	2	585	5	592
Wild/wild	2	7	1195	1204
Total	247	593	1203	2043

^a Analysis has been restricted to only the C282Y mutation.

the results of the 2043 individual sample challenges made over the initial five years of the survey (restricted to the C282Y mutation). Data were collected from the published tables and accompanying comments in the semiannual ACMG/CAP MGL Participant Summary Reports.⁵ A complete list of the sample challenges, the types of errors, and adjustments made during the analysis is available via the ArticlePlus feature at the *Genetics in Medicine* Web site (<http://www.geneticsinmedicine.org>). Overall, there are 20 errors in C282Y genotyping, for an error rate of 1.0% (95% CI 0.6%–1.5%). However, only 8 of the 20 errors involve C282Y homozygosity (four false positive homozygotes and four false negative homozygotes). Based on these data, the estimated analytic sensitivity is 98.4% (243 of 247 true homozygous sample challenges, 95% CI 95.9%–99.5%). The corresponding estimate for analytic specificity is 99.8% (1792 of 1796 true nonhomozygous sample challenges, 95% CI 99.4%–99.9%). The analytic specificity is similar for the two underlying true genotypes (C282Y/wild or wild/wild). There are too few observations to determine whether these rates vary over the five years.

The analytic positive and negative predictive powers can be computed for a hypothetical population with a C282Y homozygosity prevalence of 40 per 10,000. Among the 40 true homozygotes, 39 ($40 \times 98.4\%$) will be correctly identified as being homozygous; one will be falsely negative. Among the 9,960 true nonhomozygotes, 9,940 ($9,960 \times 99.8\%$) will be correctly identified as being nonhomozygous; 20 will be falsely positive. Thus, the analytic positive predictive power is 66% ($39/[39 + 20]$, 95% CI 39%–80%, based on the CIs for analytic specificity). If testing were to be performed in a population where C282Y homozygosity is less common (e.g., African Americans or Asian Americans), the analytic positive predictive power would be lower. The negative predictive power is very high at 99.99% ($9940/9940 + 1$), partly because homozygosity in the general population is rare.

Even with the high analytic performance estimates found in this study, a significant proportion of those identified as being homozygous for the C282Y mutation as part of routine screening in a general “low risk” population will be false-positives. It is possible that many of the false-positives are due to pre- or postanalytic errors rather than the analytic process itself. If this were true, confirmatory testing utilizing a new sample would likely correct many of them. For example, were this type of confirmatory testing to correct 90% of the false-positive test results, while maintaining the same analytic sensitivity, the positive predictive value would rise from 66% to 95% ($38/[38 + 2]$), with one additional true homozygote being incorrectly reclassified (false-negative). There is some evidence in the literature that confirmatory testing using a different technology may identify some false-positive homozygous results occurring in the analytic phase of testing.⁷

DISCUSSION

The preceding analyses do not include the H63D mutation test results that about 90% of participating laboratories rou-

tinely report. Testing for both mutations should not, in theory, adversely affect analytic performance estimates for the C282Y mutation, but a remote possibility exists that a laboratory test focusing on just a single mutation might perform better. This is not supported by the survey findings, where some genotyping errors were made among the five to nine laboratories that test only for the C282Y mutation. There is a mixture of clinical and nonclinical laboratories participating in the survey, and the nonclinical laboratories might be responsible for many of the errors. This also is not supported by survey results. In 2002, for example, there were 13 genotyping errors (six of which involved the C282Y mutation), and all were made in laboratories reporting clinical results. Errors in the survey are also not restricted to specific methods. Over the five years, errors were reported for several of the analytic methods. Sample mix-up is a likely candidate for causing some of the errors; on occasion, the correct genotypes were reported, but not in the correct order. Also, the CAP/ACMG committee has unpublished data indicating that a high proportion of genotyping errors occurring in the factor V Leiden survey are due to sample mix-ups and other clerical errors (Wayne W. Grody, personal communication, 2003). However, none of the four false-positive homozygous C282Y results occurred when a true homozygous sample was included in an ACMG/CAP distribution.

As a further consideration, a few laboratories might be responsible for a majority of errors. An external proficiency testing program for cystic fibrosis has reported that less than half the participating laboratories were error free over a three year time period.⁸ This tendency is confirmed in the ACMG/CAP data. For example, In the 2001 MGL-A Participant Summary Report, the seven incorrect responses for HFE mutation testing were reported to be from seven different laboratories (six of the seven were clinical laboratories). A European proficiency testing program for cystic fibrosis reported that one source of error occurred because the proficiency testing sample was prepared in a way that was different from the protocol routinely used by some laboratories.⁹ For example, the proficiency testing preparation (purified DNA) might be similar to that used in a participating laboratory dealing with blood samples, but might be too concentrated for another laboratory that routinely deals with buccal samples (DNA lysate). This is unlikely to be the cause of laboratory errors in the MGL survey of HFE mutation testing, however, because virtually all laboratories perform this test on blood rather than buccal samples. Lastly, it has been reported that an *HFE* primer frequently used by testing laboratories could, in the presence of a common polymorphism, result in a false-positive homozygous result in a true heterozygote. However, the ACMG/CAP Survey found that 67 U.S. laboratories (many using that primer) had all correctly genotyped a C282Y heterozygote sample that also carried the polymorphism.¹⁰ This situation is unlikely to be the cause of any of the false-positive results observed in our study.

It is difficult to compare analytic performance of different DNA tests, often because of the setting and the purpose of testing. The analytic validity of *CFTR* testing in the setting of prenatal screening for cystic fibrosis has been published.⁶ For

example, consider the identification of non-Hispanic Caucasians carriers of a cystic fibrosis mutation. Even though the analytic sensitivity and specificity for *CFTR* testing are slightly lower than for *HFE* testing reported in the present study, the high carrier rate for cystic fibrosis (1 in 25) results in a similar analytic positive predictive power (75%) to that found for C282Y homozygosity (66%). There are common lessons to be learned. Both testing methodologies are highly sensitive and specific, but not perfect. These findings suggest that confirmatory testing may be worthwhile for positive DNA screening tests in general. Our findings in a screening setting may not generalize to a diagnostic setting. For example, *HFE* testing for individuals with clinical symptoms of hemochromatosis will likely include testing for multiple mutations and genotypes other than homozygosity for C282Y will be of interest.

Population screening for *HFE* mutations is not currently recommended. The estimates of analytic performance presented in this study, however, can serve as a guidepost for those currently performing population-based *HFE* testing on a research basis,^{11–13} and also have implications for possible future testing. A proportion of individuals identified as being homozygous for the C282Y mutations might, in fact, be analytic false-positives. We propose that this proportion is likely to be highly dependent on whether confirmatory testing of homozygous test results utilizing a new sample is performed. The possibility of false-positive test results should also be considered when evaluating the published literature, especially population-based trials,^{11–13} and may be a contributing factor to the relatively low penetrance estimates for this genotype. Confirmatory testing to identify false-positive test results is likely to be an important and necessary component of any population-based screening program for *HFE* mutations.

ACKNOWLEDGMENTS

This work was performed as part of an ACCE Review and support was provided by a cooperative agreement (UR3/CCU319352) with the Centers for Disease Control and Prevention, Office of Genomics and Disease Prevention. The data source for this analysis was designed, collected, and reported by the Biochemical and Molecular Genetics Resource Committee of the College of American Pathologists and the American College of Medical Genetics. We thank the Committee members for their comments and suggestions.

References

1. Secretary's Advisory Committee on Genetic Testing. Enhancing the oversight of genetic tests: Recommendations of the SAGCT. National Institutes of Health, July 2000. Available at: <http://www4.od.nih.gov/oba/sacgt.htm>. Accessed October 6, 2003.
2. Feder JN, Gnirke A, Thomas W, Tsuchihashi Z, Ruddy DA, Basava A et al. A novel MCH class I-like gene is in patients with hereditary haemochromatosis. *Nat Genet* 1996;13:399–408.
3. Beutler E, Gelbart T, West C, Lee P, Adams M, Blacksonne R et al. Mutation analysis in hereditary hemochromatosis. *Blood Cell Mol Dis* 1996;22:187–194.
4. Haddow JE, Palomaki GE, McClain M, Craig W. Hereditary haemochromatosis and hepatocellular carcinoma in males: a strategy for estimating the potential for primary prevention. *J Med Screen* 2003;10:11–13.
5. ACMG/CAP Molecular Genetics Survey Sets, 1999, 2000, 2001, 2002. Northfield, Ill: College of American Pathologists, 1998.

6. Palomaki GE, Bradley LA, Richards CS, Haddow JE. Analytic validity of cystic fibrosis testing: A preliminary estimate. *Genet Med* 2003;5:15–20.
7. Beutler E, Gelbart T. Large-scale screening for HFE mutations: Methodology and cost. *Genet Test* 2000;4:131–142.
8. Dequeker E, Cassiman JJ. Genetic testing and quality control in diagnostic laboratories. *Nat Genet* 2000;25:259–260.
9. Cuppens H, Cassiman JJ. A quality control study of *CFTR* mutation screening in 40 different European laboratories. *Eur J Hum Genet* 1995;3:235–245.
10. Noll WW, Belloni DR, Stenzel TT, Grody WW. Polymorphism in intron 4 of *HFE* does not compromise haemochromatosis mutation results. *Nat Genet* 1999;23:271–272.
11. Beutler E, Fellitti V, Gelbart T, Ho N. The effect of *HFE* genotypes on measurements of iron overload in patients attending a health appraisal clinic. *Ann Intern Med* 2000;133:329–337.
12. Beutler E, Fellitti VJH, Loziol JA, Ho NJ, Gelbart T. Penetrance of 845G>A (C282Y) *HFE* hereditary haemochromatosis mutation in the USA. *Lancet* 2002;359:211–218.
13. Harris EL, Mellen BG, Adams PC, Acton R, Barton JC, McLaren GD et al. Hemochromatosis, and Iron Overload Screening (HEIRS) study: Prevalence of elevated iron test levels by *HFE* genotype [abstract]. *Am J Hum Genet* 2002;10:a83.