Letters to the Editor

Am. J. Hum. Genet. 67:1033–1035, 2000

Genetic Testing Should Not Be Advocated as a Diagnostic Tool in Familial Forms of Dementia

To the Editor:

In their article in the January 2000 issue of the *Journal,* Finckh et al. advocate the use of a molecular diagnostic program in patients with early-onset dementia (EOD) and a family history of dementia (Finckh et al. 2000). In 36 patients with EOD, Finckh et al. screened for mutations in the presenilin (PSEN)–1 and -2 genes (MIM 104311 and MIM 600759, respectively), the amyloid-precursor protein (APP) gene (MIM 104760), and the prion protein (PRNP) gene (MIM 176640); in 12 patients, they found mutations that were considered to be disease causing. Finckh et al. argue that, in the absence of specific antemortem diagnostic markers for familial Alzheimer disease (AD) or hereditary prion disease, molecular testing is important to ensure that treatable dementias are not missed. Although the findings by Finckh et al. are of interest, we think that the implications of these findings for clinical practice are seriously limited. We offer several arguments against the use of genetic testing as a diagnostic tool for the differential diagnosis of dementia in general practice.

One significant limit on the usefulness of genetic testing arises from the distribution and prevalence of APP, PSEN, and PRNP mutations in EOD. The usefulness of a clinical diagnostic test is determined, in large part, by the composition of the patient population. Among the disease-causing mutations found by Finckh et al., two-thirds were associated with AD, and one-third occurred in the PRNP gene. This distribution is surprising and does not reflect the typical clinical experience. The prevalence of early-onset AD is estimated to be 18.2–41.2 persons per 100,000 at risk, and that of autosomal-dominant early-onset AD is estimated to be 5.3 persons per 100,000 at risk, whereas the familial forms of the prion diseases, which include familial Creutzfeldt-Jakob disease, Gerstmann-Sträussler-Scheinkers syndrome, and fatal familial insomnia, have an estimated incidence of only 1/10 million/year (Haywood 1997; Campion et al. 1999). Even in highly specialized neurological or geriatric centers, a clinician will not encounter a demented population in which 33% of patients have prion disease. The rarity of patients with prion disease underscores the very atypical composition of the population studied by Finckh et al. and the limited relevance that their study has for ordinary diagnostic practice. Furthermore, we and others have found that, in a population-based sample, only ~20% of patients with early-onset familial AD have a causative mutation (Cruts et al. 1998; Kamimura et al. 1998), whereas Finckh et al. observed a mutation in 45% of such patients. We argue that, for clinical practice, the findings in a population-based study are more relevant than those in a highly selected population.

Our second concern is related to the application of molecular screening of dementia genes in clinical practice. In contrast to the mutations in the APP gene, which are clustered around exons 16 and 17, a large number of rare mutations in the PSEN genes are known to be distributed throughout the gene. More than 50% of these PSEN mutations are genetically "private"; that is, they are found only in a particular patient or family (Blacker and Tanzi 1998; Cruts et al. 1998). As shown by Finckh et al., novel mutations are still found. Also, previously unknown mutations are detected in the PRNP gene (Laplanche et al. 1999). The causative effects of these are difficult to interpret. A notorious example of misjudging the pathogenicity of a presumed missense mutation is the Glu318Gly substitution in the PSEN-1 gene. For example, an 86-year-old woman fulfilling the National Institute of Neurological Disorders and Stroke and the Alzheimer Disease and Related Disorders Association criteria for probable AD was referred to the Memory Clinic of Erasmus University Medical School. Her family history revealed early- and late-onset AD in several first-degree relatives. We therefore screened for the known AD-associated genes, and we detected an E318G mutation in the PSEN-1 gene. This substitution was earlier reported as a causative mutation in patients with familial early-onset AD (Cruts and Van Broeckhoven 1998). In 1999, however, Dermaut et al. demonstrated that an elderly group of 256 control subjects included 9 carriers of this substitution who were not demented, results indicating that the frequency in control sub-

jects was similar to that found in patients (Dermaut et al. 1999). Rather than being a pathogenic mutation, the E318G is a rare allele, that is not associated with either AD or dementia in general and does not influence the β-amyloid formation. In the absence of any population data, we might have incorrectly reasoned that our patient represented a genetic case with late onset. This example illustrates that, even in familial cases, studies of a large series of controls should be performed before conclusions are drawn about the pathogenicity and penetrance of a particular mutation. For any untreatable disease with a devastating course, as is the case in AD and prion diseases, the burden of an incorrect molecular diagnosis should be prevented by all possible means, since genetic testing does not have implications for the patient alone but also discloses predictive information to family members, which could influence such issues as life expectancy, insurability, and psychosocial well-being.

We are especially concerned by the emphasis given to the use of molecular screening of AD genes and the PRNP gene, in light of the importance of ascertaining the presence of treatable dementias. First, considerably easier ways to diagnose treatable dementias exist. In clinical settings it is more straightforward to test for the presence of a treatable dementia directly (e.g., by measuring levels of thyroid-stimulating hormone level to test for dementia associated with hypothyroidism). Second, the presence of a mutation does not eliminate the possibility of the coexistence of a treatable form of dementia. Conditions such as depression, drug intoxication, vascular dementia, and metabolic disorders can mimic and coincide with AD, especially in patients with a long disease course. Even when a major mutation is present in the family, tests for treatable causes of dementia should not be omitted in clinical practice.

Finckh et al. report interesting data on novel mutations. Furthermore, they raise an intriguing question regarding the use of genetic testing as a diagnostic tool. However, it is of major importance to recognize their report's limited applicability to clinical practice. We argue that the contribution of genetic testing to clinical diagnosis is small and does not counterbalance the problems associated either with interpretation of any mutation that is found or with secondary effects on family members. Nevertheless, for scientific reasons, genetic testing is very worthwhile. Testing may increase our knowledge about the different mutations, which could have clinical applications in the future. However, the limits of current knowledge are too great to justify clinical use of genetic testing in the diagnostic process.

Esther A. Croes,[1] Bart Dermaut,[3] Tischa J. M. van der Cammen,[2] Christine van Broeckhoven,[3] and Cornelia M. van Duijn[1]

[1]*Department of Epidemiology and Biostatistics and* [2]*Memory Clinic, Department of Internal and Geriatric Medicine, Erasmus University Medical School, Rotterdam; and* [3]*Laboratory of Molecular Genetics, Flanders Interuniversity Institute of Biotechnology, Born-Bunge Foundation, University of Antwerp, Antwerp*

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for PSEN-1 and -2 [MIM 104311 and MIM 600759, respectively], APP [MIM 104760], and PRNP [MIM 176640])

References

Blacker D, Tanzi RE (1998) The genetics of Alzheimer disease: current status and future prospects. Arch Neurol 55:294–296

Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, Thomas-Anterion C, Michon A, Martin C, Charbonnier F, Raux G, Camuzat A, Penet C, Mesnage V, Martinez M, Clerget-Darpoux F, Brice A, Frebourg T (1999) Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. Am J Hum Genet 65:664–670

Cruts M, Van Broeckhoven C (1998) Presenilin mutations in Alzheimer's disease. Hum Mutat 11:183–190

Cruts M, van Duijn CM, Backhovens H, Van den Broeck M, Wehnert A, Serneels S, Sherrington R, Hutton M, Hardy J, St George-Hyslop P, Hofman A, Van Broeckhoven C (1998) Estimation of the genetic contribution of presenilin-1 and -2 mutations in a population based study of presenile Alzheimer disease. Hum Mol Genet 7:43–51

Dermaut B, Cruts M, Slooter AJ, Van Gestel S, De Jonghe C, Vanderstichele H, Vanmechelen E, Breteler MB, Hofman A, van Duijn C, Van Broeckhoven C (1999) The Glu318Gly substitution in presenilin 1 is not causally related to Alzheimer disease. Am J Hum Genet 64: 290–292

Finckh U, Muller-Thomsen T, Mann U, Eggers C, Marksteiner J, Meins W, Binetti G, Alberici A, Hock C, Nitsch RM, Gal A (2000) High prevalence of pathogenic mutations in patients with early-onset dementia detected by sequence analyses of four different genes. Am J Hum Genet 66:110–117

Haywood AM (1997) Transmissible spongiform encephalopathies. N Engl J Med 337:1821–1828

Kamimura K, Tanahashi H, Yamanaka H, Takahashi K,

Asada T, Tabira T (1998) Familial Alzheimer's disease genes in Japanese. J Neurol Sci 160:76–81

Laplanche JL, Hachimi KH, Durieux I, Thuillet P, Defebvre L, Delasnerie-Laupretre N, Peoc'h K, Foncin JF, Destee A (1999) Prominent psychiatric features and early onset in an inherited prion disease with a new insertional mutation in the prion protein gene. Brain 122:2375–2386

Address for correspondence and reprints: Dr. C. M. van Duijn, Department of Epidemiology and Biostatistics, Erasmus University Medical School Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. E-mail: duijn@epib.fgg.eur.nl

0002-9297/2000/6704-0032$02.00

Am. J. Hum. Genet. 67:1035–1036, 2000

Reply to Croes et al.

To the Editor:

Croes et al. (2000 [in this issue]) make the point that genetic testing as a diagnostic tool shows poor performance in differential diagnosis in general medical practice. We fully agree with this comment. Therefore, one of the goals of our study (Finckh et al. 2000) was to establish criteria that would increase the chance of identifying a pathogenic mutation in the setting of a specialized clinic. Indeed, among patients who had both onset at an early age and positive family history for early-onset dementia (EOD), diagnostic sequencing identified disease-relevant mutations in >50% of the patients analyzed by us. Another notable result of our study was the finding of four prion mutations among the 36 EOD patients, which suggested that atypical forms of prion disease may remain underdiagnosed. This assumption is supported by independent observations, such as those made by two coauthors of the letter by Croes et al. (2000), who found a *PRNP* insertion mutation in a patient with both prion disease and ante mortem diagnosis of familial Alzheimer disease (FAD) (Dermaut et al. 1998).

We agree with Croes et al. that assessment of the relevance of previously unknown mutations is a difficult issue. Nonetheless, in recent screening studies of FAD, 72%–83% of the sequence changes corresponded to pathogenic mutations already reported (Kamimura et al. 1998; Campion et al. 1999). In our study, 58% of the mutations had been previously described by others. Repeated identification of any given rare mutation in a rare disorder, together with the absence of the mutation in control probands, significantly increases the likelihood that it has causative effects.

We were pleased to see that Croes et al. agree with our conclusion that E318G in PS1 is a nonpathogenic polymorphism and that they reemphasize the importance of a careful and critical analysis of the literature. The importance of early and disease-specific diagnosis of EOD as a way of identifying treatable forms of dementia is an issue separate from our assertion that diagnostic sequencing of the four known EOD genes may provide important information for proper clinical and genetic counseling in the early phases of the disorder.

Ulrich Finckh,[1] Tomas Müller-Thomsen,[2] Ulrike Mann,[2] Christian Eggers,[3] Josef Marksteiner,[5] Wolfgang Meins,[4] Giuliano Binetti,[6] Antonella Alberici,[6] Christoph Hock,[7,8] Roger M. Nitsch[8] and Andreas Gal[1]

Departments of [1]Human Genetics, [2]Psychiatry, and [3]Neurology, University Hospital Eppendorf, University of Hamburg, and [4]Albertinen-Haus, Geriatric Centre, Hamburg; [5]Clinic of Psychiatry, University of Innsbruck, Innsbruck; [6]IRCCS Centro S. Giovanni di Dio, Alzheimer Disease Unit, Brescia, Italy; [7]Department of Psychiatry, University of Basel, Basel, and [8]Department of Psychiatry Research, University of Zurich, Zurich

References

Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, Thomas-Anterion C, Michon A, Martin C, Charbonnier F, Raux G, Camuzat A, Penet C, Mesnage V, Martinez M, Clerget-Darpoux F, Brice A, Frebourg T (1999) Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. Am J Hum Genet 65:664–670

Croes EA, Dermaut B, van der Cammen C, van Broeckhove C, van Duijn CM (2000) Genetic testing should not be advocated as a diagnostic tool in familial forms of dementia. 67:000–000

Dermaut B, De Jonghe C, Cras P, Backhovens H, Slooter AJC, Hofman A, Breteler MM, van Duijn CM, Hendriks L, van Broeckhoven C, Cruts M (1998) Variable expression of the Glu318Gly mutation in presenilin-1 suggests reduced penetrance. Paper presented at the Foundation IPSEN conference, Epidemiolgie de la Maladie d'Alzheimer: Du Gène à la Prévention. Paris, May 25

Finckh U, Müller-Thomsen T, Mann U, Eggers C, Marksteiner J, Meins W, Binetti G, Alberici A, Hock C, Nitsch RM, Gal A (2000) High prevalence of pathogenic mutations in patients with early-onset dementia detected by sequence analyses of four different genes. Am J Hum Genet 66:110–117

Kamimura K, Tanahashi H, Yamanaka H, Takahashi K, Asada T, Tabira T (1998) Familial Alzheimer's disease genes in Japanese. J Neurol Sci 160:76–81

Address for correspondence and reprints: Dr. Ulrich Finckh, Universitätsklinikum Hamburg-Eppendorf, Institut für Humangenetik, Butenfeld 42, 22529 Hamburg, Germany. E-mail: finckh@uke.uni-hamburg.de

0002-9297/2000/6704-0033$02.00

Am. J. Hum. Genet. 67:1036–1039, 2000

The Efficiency of Pooling in the Detection of Rare Mutations

To the Editor:

After citing a variety of uses of pooled testing in genetic studies, Amos et al. (2000) suggested that mutations in individual patients could be detected more efficiently by being tested in pools. A typical mutation-detection protocol requires that many segments of the gene—for example, an amplicon consisting of one or a few close exons—need to be evaluated for detection of a mutation. Thus, even if the mutation has a prevalence of ~2%, as in the case of *BRCA1* or *BRCA2* in Ashkenazim (Hartge et al. 1999), the probability that any segment will contain a mutation is much smaller, perhaps on the order of .0005–.005. The use of pools or groups of samples to identify individuals or to estimate the prevalence of such a rare characteristic has been extensively studied in the statistical literature (Dorfman 1943; Sobel and Elashoff 1972; Gastwirth and Hammick 1989; Tu et al. 1995; Brookmeyer 1999). Using the corrected formula (see the erratum by Amos et al. [in this issue]) for the number of runs or tests needed to identify individuals with a mutation, one can fully appreciate the potential of pooling methods. A variant of the grouping procedure is described that in some circumstances leads to greater gains in efficiency when grouped testing is utilized.

The sensitivity of an assay—that is, the probability that a mutation will be detected, given that at least one member of the pool has it—is a potential limiting factor in practice. For screening of individuals to determine their carrier status, the sensitivity should be as close as possible to 100%. For detection of mutations by multiplex single-nucleotide primer extension, 100% sensitivity was achieved in pools of size 10–20 but dropped to 80% in pools of 30 (Krook et al. 1992). When denaturing high-performance liquid chromatography was used to identify *BRCA* mutations, 100% sensitivity was observed for several amplicons studied in groups of size five to nine (J. Rutter, personal communication). Thus, for the largest pool size for which a mutation detector is 100% sensitive, it is helpful to know the largest mutation prevalence for which pooling is efficient.

Suppose that the prevalence of a mutation in a single unit (exon or amplicon) being studied is π and that n individuals donate samples. For pools of size r, the probability, γ, that at least one member of the pool has a mutation is $1 - (1 - \pi)^r$. Assume that the test is 100% accurate in classifying a pool as having or not having a mutation. Since Y, the number of runs or tests that need to be done without pooling is n, for any pooling protocol in which the ratio of the expected value (y) of $Y{:}n < 1$, the strategy saves runs. We denote this ratio by F, for fraction of tests required relative to individual testing; and the efficiency of a pooling method is $1 - F$, the fraction of tests saved. When the classical single-stage pooling method (Dorfman 1943), which retests, one at a time, the individuals in a positive pool, is used, the expected number of runs needed to completely identify all the mutations in the segment under study in the sample of n individuals is

$$E(Y) = \left(\frac{n}{r}\right) + n\gamma \ . \qquad (1)$$

The derivation follows. The probability that a pool contains a mutation, which implies that it will test positive, is γ. Since all r individuals in the pool will be tested, a positive pool receives a total of $r + 1$ tests. The probability that a pool is negative is $(1 - \gamma)$. Those pools are classified with one test, so the expected number of tests per pool is $(r + 1)\gamma + (1 - \gamma) = 1 + \gamma r$. Since there are $\frac{n}{r}$ pools, the expected number of tests is given by equation (1). Note that the prevalence, π, enters into equation (1) because it determines the probability, γ, that a pool is positive.

Amos et al. (2000) also considered the situation in which there is a probability β, of a false-positive result in a pool—that is, $1 - \beta$ is the specificity of the mutation-detection process while the sensitivity remains perfect. The same reasoning that led to equation (1) shows that the expected number, y, of runs or tests is given by

$$y = n\left\{\frac{1}{r} + [1 - (1 - \beta)(1 - \pi)^r]\right\} . \qquad (2)$$

From equations (1) and (2), we can calculate the range of values of π for which the ratio of the expected number, y, of tests or runs (Y) to the total sample size, n, is <1, which implies that pooling is at least as efficient as individual testing. We also present the largest π value, $\pi_{.5}$, for which $\frac{y}{n} < .5$, which indicates that pooling will result in a substantial savings in the ex-

Table 1

Mutation Prevalence for Which Pooling Is Efficient, as a Function of Pool Size

Pool Size	Largest Prevalence for Which $\frac{y}{n}$ Is <1		<.5	
	$\beta = 0$	$\beta = .05$	$\beta = 0$	$\beta = .05$
2	.293	.275	Not possible	
3	.307	.295	.059	.043
4	.293	.284	.069	.057
5	.275	.268	.069	.059
7	.243	.237	.061	.054
10	.206	.202	.050	.045
12	.187	.184	.044	.040
15	.165	.162	.037	.034
20	.139	.137	.029	.027
25	.121	.119	.024	.022
40	.087	.085	.016	.014
50	.075	.074	.013	.012
75	.056	.055	.009	.008
100	.045	.045	.007	.006

pected number of tests. For the case of perfect tests, these values of π_1 and $\pi_{.5}$ are given by

$$\pi_1 \leqslant 1 - \left(\frac{1}{r}\right)^{\frac{1}{r}} \text{ and } \pi_{.5} \leqslant 1 - \left(\frac{2+r}{2r}\right)^{\frac{1}{r}} . \qquad (3)$$

When the specificity is $1 - \beta$, the equations become

$$\pi_1 \leqslant 1 - \left[\frac{1}{r(1-\beta)}\right]^{\frac{1}{r}} \text{ and } \pi_{.5} \leqslant 1 - \left(\frac{.5 + \frac{1}{r}}{1-\beta}\right)^{\frac{1}{r}} . \qquad (4)$$

In table 1, I present the values of π_1 and $\pi_{.5}$ that are obtained from equations (3) and (4), as a function of r, the pool size. The results for π_1 indicate that pooling in relatively small pools, up to size five or six, can be efficient for values of $\pi \leqslant .25$. Moreover, pools of size $\leqslant 10$ can save at least half of the runs, for prevalences $\leqslant .045$, even with a 5% false-positive rate. Indeed, a small lack of specificity does not have a major impact on the range of prevalences for which pooling is useful. For the exons and amplicons occurring in DNA mutation research, in which the prevalence of a mutation at a specific segment being examined is likely to be near .001, pools of 40–100 individual samples would be quite efficient. Of course, this assumes that the sensitivity of the test remains perfect in such samples. Thus, the major limitation in the use of pooling techniques is the maximum size of the group for which the sensitivity of the test is essentially 1.

For a specific prevalence π, the optimum pool size is obtained by differentiating equations (1) and (2), respectively, and by setting the derivative to 0. When the test used has perfect sensitivity and specificity, r satisfies

$$r \ln(1-\pi) + \ln\ln\left(\frac{1}{1-\pi}\right) = -2\ln r \ ;$$

when the specificity is $1-\beta$, the optimum pool size r satisfies

$$\ln(1-\beta) + r\ln(1-\pi) + \ln\ln\left(\frac{1}{1-\pi}\right) = -2\ln r \ .$$

The values of r that yield the optimum pool size for a range of prevalences is given in table 2. A small false-positive rate ($\beta = .05$) does not have a noticeable impact on the optimal group size but does diminish the efficiency gain in the very-small-prevalence setting when large pools are possible. The results in table 2 indicate that pooling strategies have a greater potential of improving the efficiency of mutation testing than previous results had indicated; for example, when $\pi = .01$, the data in table 2 indicate that, for the Dorfman procedure, the optimal pool size is 11 and the expected number of tests is 20%–24% of the number, n, of individuals, depending on whether $\beta = 0$ or .05.

Although a complex multistage sampling protocol may not be appropriate when the optimal pool size is <10 (Amos et al. 2000), a one-step procedure can improve the efficiency of grouping. Consider a pool size $r = 2m$. If the pool tests negative, then all units are mutation free. When a pool tests positive, it is divided into two pools of size m that are tested. For rare mutations, usually only one of the two pools will be positive, so only m further tests are needed. A simple upper bound, y_u, for the expected number, y, of tests used by this one-step method is obtained by assuming that, in a positive pool, if there are two or

Table 2

Optimal Pooling Size and Fraction of Tests Required, Relative to Individual Testing, for Two Pooling Methods

π	Optimal Pool Size (% of Tests Required) $\beta = 0$ Dorfman	$\beta = 0$ One Step	$\beta = .05$ Dorfman	$\beta = .05$ One Step
.2	3 (82.1)	4 (93.1)	3 (84.7)	4 (94.7)
.1	4 (59.4)	5 (60.9)	4 (62.7)	5 (62.6)
.05	5 (42.6)	8 (40.6)	5 (46.5)	7 (41.6)
.01	11 (19.6)	14 (16.0)	11 (24.0)	15 (16.9)
005	15 (13.9)	20 (11.0)	15 (18.5)	21 (11.6)
.001	32 (6.3)	45 (4.7)	33 (11.1)	47 (5.1)
.0005	45 (4.5)	63 (3.3)	46 (9.3)	66 (3.6)
.0001	100 (2.0)	142 (1.4)	103 (6.9)	149 (1.6)

more positive individuals, both half-groups will test positive and all r units will need to be tested. When the false-positive rate for testing a group of size r is β, it is reasonable to assume that the error rate of the test for a pool with half as many individuals ($\frac{r}{2}$) is $\frac{\beta}{2}$. Denote the probability that a pool has exactly one positive individual by $\eta = r\pi(1 - \pi)^{r-1}$. In this case, the expected number of tests for a pool is

$$1 + \gamma(2 + r) + \eta\left(\frac{\beta}{2} - 1\right)\frac{r}{2} + (1 - \gamma)\beta\left(2 + \frac{r\beta}{2}\right). \quad (5)$$

The fraction, F, of tests needed relative to individual testing is $\frac{1}{r}$ times equation (5). When $\beta = 0$, the upper bound for F for the one-step method becomes

$$y_u = \left(\frac{n}{r}\right)\left[1 + \gamma(2 + r) - \frac{\eta r}{2}\right].$$

The optimal pool size and fraction of tests with regard to the size, n, of the population screened, as required by the Dorfman and one-step procedures, are given in table 2. When the tests are perfect, the one-step procedure does not yield a substantial increase in efficiency until fairly large pools of a very-low-prevalence mutation can be pooled. The one-step method provides efficiency gains over a larger range of prevalence values and modest pool sizes when there are false-positive pools. This occurs because those pools are truly negative and because there is a very high probability that the two half-pools will be classified correctly by the two tests.

The results indicate that pooling should be quite helpful when a large population is being screened for relatively rare genetic mutations, especially when the prevalence of a mutation in an exon or amplicon is likely to be <.001. As improved technology enables larger pools to be examined (Zarbl et al. 1998), the efficiency of the one-step method should reduce the costs substantially; for example, if the prevalence is .005 and 20 individual samples can be pooled, the number of tests needed is only ~11% of the number of individuals screened. Greater savings can be achieved, at low prevalences, with multistage (Brookmeyer 1999) or repooling (Munoz-Zanzi et al. 2000) plans.

The formulas for the optimum pool size depend on the prevalence of the mutation in the amplicon assayed. Since this may not be known precisely, one can adopt a two-stage procedure (Hughes-Oliver and Swallow 1994) that changes the pool size on the basis of the estimated prevalence for a partial sample. The results in table 2 can be used to determine the group size for the remaining analyses.

There are several other potential applications of pooling to mutation detection. The methods discussed both here and by Amos et al. (2000) assume perfect sensitivity. In practice, errors occur, so it is useful to use pooling methods to retest a sample of the screened negatives, both to confirm that the sensitivity remains essentially perfect and to ensure that individuals with the mutation are not missed. Such a procedure has been shown to be a cost-effective quality-control method for blood screening (Gastwirth and Johnson 1994). Group testing, without identification of individuals, has also been shown to yield accurate estimates of the prevalence of a rare disease or trait (Gastwirth and Hammick 1989), while preserving the privacy of participants.

Acknowledgments

It is a pleasure to thank Mitchell Gail, Joni Rutter, and Binbing Yu for helpful discussions.

JOSEPH L. GASTWIRTH
Division of Cancer Epidemiology and Statistics, National Cancer Institute and Department of Statistics, George Washington University, Washington, DC

References

Amos CI, Frazier ML, Wang W (2000) DNA pooling in mutation detection with reference to sequence analysis. Am J Hum Genet 66:1689–1692

Brookmeyer R (1999) Analysis of multistage pooling studies of biological specimens for estimating disease incidence and prevalence. Biometrics 55:608–612

Dorfman R (1943) The detection of defective members of large populations. Ann Math Stat 14:436–440

Gastwirth JL, Hammick P (1989) Estimation of the prevalence of a rare disease preserving the anonymity of the subjects by group testing: application to estimating the prevalence of AIDS antibodies in blood donors. J Stat Planning Inference 22:15–27

Gastwirth JL, Johnson WO (1994) Screening with cost-effective quality control: potential applications to HIV and drug testing. J Am Stat Assoc 89:972–981

Hartge P, Struewing JP, Wacholder S, Brody LC, Tucker MA (1999) The prevalence of common BRCA1 and BRCA2 mutations among Ashkenazi Jews. Am J Hum Genet 64: 963–970

Hughes-Oliver JM, Swallow WH (1994) A two-stage adaptive group-testing procedure for estimating small proportions. J. Am Stat Assoc 89:982–993

Krook A, Stratton IM, O'Rahilly S (1992) Rapid and simultaneous detection of multiple mutations by pooled and multiplex single nucleotide primer extension: application to the study of insulin-responsive glucose transporter and insulin receptor mutations in non-insulin-dependent diabetes. Hum Mol Genet 1:391–395

Munoz-Zanzi CA, Johnson WO, Thurmond MC, Hietala

SK (2000) Pooled-sample testing as a herd screening tool for detection of bovine viral diarrhea virus in persistently infected cattle. J Vet Diagn Invest 12:195–203

Sobel M, Elashoff RM (1975) Group testing with a new goal, estimation. Biometrika 63:181–193

Tu XM, Litvak E, Pagano M (1995) On the informativeness and accuracy of pooled testing in estimating prevalence of a rare disease: application to HIV screening. Biometrika 82:287–297

Zarbl H, Aragaki C, Zhao LP (1998) An efficient protocol for rare mutation genotyping in a large population. Genet Test 2:315–321

Address for correspondence and reprints: Dr. Joseph L. Gastwirth, Biostatistics Branch, Division of Cancer Genetics and Epidemiology, National Cancer Institute, 6120 Executive Boulevard, MS 7244, Bethesda, MD 20892. E-mail: jlgast@research.circ.gwu.edu

0002-9297/2000/6704-0034$02.00