



ELSEVIER

Journal of Statistical Planning and
Inference 96 (2001) 167–177

Journal of
statistical planning
and inference

www.elsevier.com/locate/jspi

Effects of violations of assumptions on likelihood methods for estimating the penetrance of an autosomal dominant mutation from kin-cohort studies

Mitchell H. Gail^{a, *}, David Pee^b, Raymond Carroll^c

^a*Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6130 Executive Boulevard, EPN 431, Rockville, MD 20892, USA*

^b*Information Management Services, Rockville, MD, USA*

^c*Texas A&M University, College Station, TX, USA*

Abstract

Struewing et al. (1997) used the kin-cohort design to estimate the risk of breast cancer in women with autosomal dominant mutations in the genes BRCA1 and BRCA2. In this design, a proband volunteers to be genotyped and then reports the disease history (phenotype) of his or her first-degree relatives. Gail et al. (1999) developed maximum likelihood estimation of parameters for autosomal dominant genes with the kin-cohort design. In this paper we examine the effects of violations of key assumptions on likelihood-based inference. Serious overestimates of disease risk (penetrance) and allele frequency result if people with affected relatives tend to volunteer to be probands more readily than people without affected relatives. Penetrance will be underestimated if probands fail to report all the disease present among their relatives, and serious overestimates of penetrance and allele frequency can result if probands give false positive reports of disease. Sources of familial disease aggregation other than the gene under study result in overestimates of the penetrance in mutation carriers, underestimates of penetrance in non-carriers, and overestimates of allele frequency. Unless sample sizes are quite large, confidence intervals based on the Wald procedure can have subnominal coverage; limited numerical studies indicate that likelihood ratio-based confidence intervals perform better. Published by Elsevier Science B.V.

Keywords: kin-cohort design; genotyped-proband design; penetrance; cancer risk; robustness; segregation analysis

1. Introduction

Much of genetic epidemiology has been concerned with inferring properties of genes that could be hypothesized, but not measured, by using segregation analysis (Elston and Stewart, 1971). The advent of techniques to identify and measure specific mutations allows one to estimate the effects of mutations on disease risk directly. Such studies are

* Corresponding author. Tel.: +1-301-496-4196; fax: +1-301-402-0081.
E-mail address: mit@cu.nih.gov (M.H. Gail).

useful for determining risk in the general population, because highly affected families used to identify genetic mutations may yield misleading estimates of disease risk.

Gail et al. (1999), hereafter referred to as GPBC, compared three population-based designs for estimating the penetrance of an autosomal dominant mutation, namely the probability that a mutation carrier would develop disease. They discussed a cohort design, in which the disease status of mutation carriers is determined by medical surveillance of the carriers, and a population-based case–control design. In the latter design, the genotype distribution of representative diseased subjects (cases) is compared with that of representative non-diseased subjects (controls), and information on the probability of disease in the population is added to estimate penetrance from Bayes' theorem (Cornfield, 1951).

The third design GPBC considered was developed by Wacholder et al. (1998), who called it the kin-cohort design. Volunteers with or without disease are genotyped, just as in a case–control design, but, in addition, one inquires about the disease status (“phenotype”) of first-degree relatives of the volunteer (the “proband”). Because the distribution of phenotypes in relatives of probands who are mutation carriers is a mixture of distributions of carriers and non-carriers, and because the distribution of phenotypes in relatives of non-carrier probands has different mixing proportions, it is possible to extract information on the underlying penetrances for carriers and non-carriers from these data (Wacholder et al., 1998). Struewing et al. (1997) used the kin-cohort design to demonstrate that mutations of the genes BRCA1 and BRCA2 carry lower risks of breast cancer than previously estimated from families with many affected members.

In the kin-cohort design, information on other members of the proband's family would ordinarily include phenotype and might, in special studies, also include genotypes of some relatives. Phenotypes could be quantitative traits, time-to-disease-onset data, or dichotomous outcomes. We restrict attention to dichotomous outcomes, although GPBC treat the parametric case of time-to-response data as well.

Special assumptions underlie a standard likelihood analysis of the kin-cohort design, as described by GPBC, who used the term genotyped proband design instead of kin-cohort design. These assumptions, some of which could be relaxed or modified, are

- (A1) risk follows an autosomal dominance pattern, in which carriers of the mutant allele have a chance (penetrance) ϕ_1 of developing disease and non-carriers have penetrance ϕ_0 ;
- (A2) the mutant allele, A , and normal allele, a , are in Hardy–Weinberg equilibrium, and mating is at random in the population;
- (A3) conditional on genotype, a relative's phenotype is independent of the phenotypes of the proband and of other relatives;
- (A4) probands are representatives of other members of the population with the same phenotype;
- (A5) disease status is determined without error; and
- (A6) sample sizes are large enough to justify standard asymptotic theory.

In this paper, we review the likelihood analysis for the kin-cohort design and investigate the effects of violations of these assumptions.

For concreteness, we assume that the proband is a mother with two daughters, as would be relevant to studies of the breast cancer genes, BRCA1 or BRCA2. Very similar results would hold for rare alleles, however, if the proband were a daughter, instead (see GPBC). Moreover, this example applies to any autosomal dominant disease for a proband and two susceptible first-degree relatives.

2. Methods

Let $g_0 = 1$ or 0 depending on whether the mother carries a mutant allele A or not. If q is the frequency of A in the population, it follows from Hardy–Weinberg equilibrium that $P(g_0 = 1) = q^2 + 2q(1 - q)$ and $P(g_0 = 0) = (1 - q)^2$. These three terms in q correspond to genotypes AA , Aa and aa , respectively. Let $g_{11} = 1$ or 0 depending on whether the first daughter carries A or not, and let $g_{12} = 1$ or 0 depending on whether the second daughter carries A or not. Let the phenotype indicators Y_0 , Y_{11} and Y_{12} be 1 or 0 depending on whether the mother or two sisters have disease or not.

The likelihood for a given family is $P(Y_{11}, Y_{12}g_0|Y_0)$, which, from Assumption (A3), reduces to

$$P(g_0|Y_0)P(Y_{11}, Y_{12}|g_0). \quad (1)$$

The term $P(g_0|Y_0)$ is used instead of $P(g_0, Y_0)$ because we are willing to assume that probands are representative, conditional on their disease status, but not unconditionally. This allows us, for example, to include all available probands with $Y_0 = 1$ but only a small fraction of those with $Y_0 = 0$. We use the term $P(Y_{11}, Y_{12}|g_0)$ instead of $P(Y_{11}, Y_{12}|g_0, Y_0)$, because under the conditional independence Assumption (A3), Y_{11} and Y_{12} depend on the proband only through the genotype of the proband.

From Bayes' theorem

$$P(g_0 = 1|Y_0 = 1) = \{q^2 + 2q(1 - q)\}\phi_1 / [\{q^2 + 2q(1 - q)\}\phi_1 + (1 - q)^2\phi_0]$$

and a similar expression can be worked out for $P(g_0 = 1|Y_0 = 0)$.

From conditional independence

$$P(Y_{11}, Y_{12}|g_0) = \sum_{g_{11}, g_{12}} P(Y_{11}|g_{11})P(Y_{12}|g_{12})P(g_{11}, g_{12}|g_0),$$

where, for example, $P(Y_{11} = 1|g_{11} = 1) = \phi_1$. The quantity $P(g_{11}, g_{12}|g_0)$, which is a function of q but not ϕ_1 or ϕ_0 , is obtained from standard Mendelian calculations as in GPBC.

Thus, the likelihood is the product over families of Eq. (1), and standard methods can be used to obtain the maximum likelihood estimates (mles) of ϕ_1 , ϕ_0 and q . GPBC used general numerical methods to maximize the log-likelihood and numerical differentiation of minus twice the log-likelihood, evaluated at the mles $\hat{\phi}_1, \hat{\phi}_0, \hat{q}$, to estimate the information matrix and hence the standard errors of $\hat{\phi}_1, \hat{\phi}_0, \hat{q}$.

We calculated 95% Wald confidence intervals for ϕ_1 , for example, as $\hat{\phi}_1 = \pm 1.96\{\text{VAR}(\hat{\phi}_1)\}^{1/2}$. One could also produce a test-based confidence interval for ϕ_1 based on the profile likelihood in ϕ_1 with limits determined by minus twice the log-profile-likelihood ratio = 3.84. Patefield (1977) describes properties of the profile likelihood, which he termed the maximized likelihood function. To determine whether this profile-likelihood confidence interval covers the true penetrance, ϕ_1 , one need only test whether minus twice the log-profile-likelihood ratio that compares $\hat{\phi}_1$ with ϕ_1 is less than 3.84.

We simulated data by selecting a mother and father at random from the general population and applying standard Mendelian genetics to generate g_0 , g_{11} , and g_{12} . Then Bernoulli values Y_0 , Y_{11} and Y_{12} were generated, conditional on g_0 , g_{11} and g_{12} , respectively. In this way, a large number of families were generated. From these families, a certain number was sampled at random without replacement from those with $Y_0 = 1$ and from those with $Y_0 = 0$. Simulations were conducted using programs written in GAUSS, Version 3.1 (1993).

Specific methods for simulations are described with corresponding results. In the standard case, however, a sample of 25,000 families with specific genotypes is generated by assuming random mating with respect to the disease gene, which is in Hardy–Weinberg equilibrium, and Mendelian transmission of genes to offspring. Then, given such genotypes and ϕ_0 and ϕ_1 , phenotypes are generated. A random sample of probands and their families is selected such that 10% of the probands are cases ($Y_0 = 1$). The total number of selected families was chosen to yield a precision on $\hat{\phi}_1$ of $1.96 \times \{\text{Var}(\hat{\phi}_1)\}^{1/2} = 0.05$, or $\pm 5\%$, as in GPBC.

3. Results

3.1. Small sample behavior

The first example is reminiscent of data for breast cancer (Claus et al., 1991) to reflect a rare autosomal dominant gene ($q = 0.01$) with high penetrance, $\phi_1 = 0.9$ for carriers, and non-negligible risk for non-carriers, $\phi_0 = 0.1$. If 10% of the probands are cases, 5893 families are required to achieve $\pm 5\%$ precision on $\hat{\phi}_1$. Note from simulation 1 in Table 1 that the estimates are unbiased and the coverage of both the Wald and likelihood ratio-based confidence intervals are near the nominal 0.95 level. Histograms of the distributions of $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} are symmetric about the population parameter values (Fig. 1). As the sample sizes are reduced, there is little evidence of bias in the estimates (simulations 2 and 3), but the coverage of the Wald confidence interval degrades seriously. Histograms corresponding to simulation 3 with 737 families in the sample indicate skewness to the right in the distributions of $\hat{\phi}_0$ and \hat{q} and a distinctly non-normal distribution of $\hat{\phi}_1$, with a point mass at the boundary $\hat{\phi}_1 = 1.0$ (Fig. 2). Indeed, $\hat{\phi}_1$ was 1.0 in 222 of 1000 simulations. It is remarkable that the coverage of the likelihood ratio confidence interval is so good in this case.

Table 1

Small sample behavior of $\hat{\phi}_1$, $\hat{\phi}_0$ and \hat{q} ^a

Simulation	True parameter values	No. probands	Design precision on $\hat{\phi}_1$ (%)	Mean estimates			Coverage of confidence intervals for ϕ_1	
				$\hat{\phi}_1$	$\hat{\phi}_0$	\hat{q}	Wald	Likelihood ratio
1	$\phi_1 = 0.9, \phi_0 = 0.1, q = 0.01$	5893	± 5	0.900	0.100	0.010	0.932	0.945
2	$\phi_1 = 0.9, \phi_0 = 0.1, q = 0.01$	1474	± 10	0.898	0.100	0.010	0.891	0.923
3	$\phi_1 = 0.9, \phi_0 = 0.1, q = 0.01$	737	± 14.1	0.905	0.100	0.010	0.773	0.942
4	$\phi_1 = 0.9, \phi_0 = 0.1, q = 0.10$	577	± 5	0.899	0.100	0.100	0.941	0.943
5	$\phi_1 = 0.5, \phi_0 = 0.1, q = 0.10$	1705	± 5	0.500	0.100	0.100	0.942	0.928

^aNumbers of repetitions in each simulation were, respectively, 5000, 1000, 1000, 4000, and 5000.

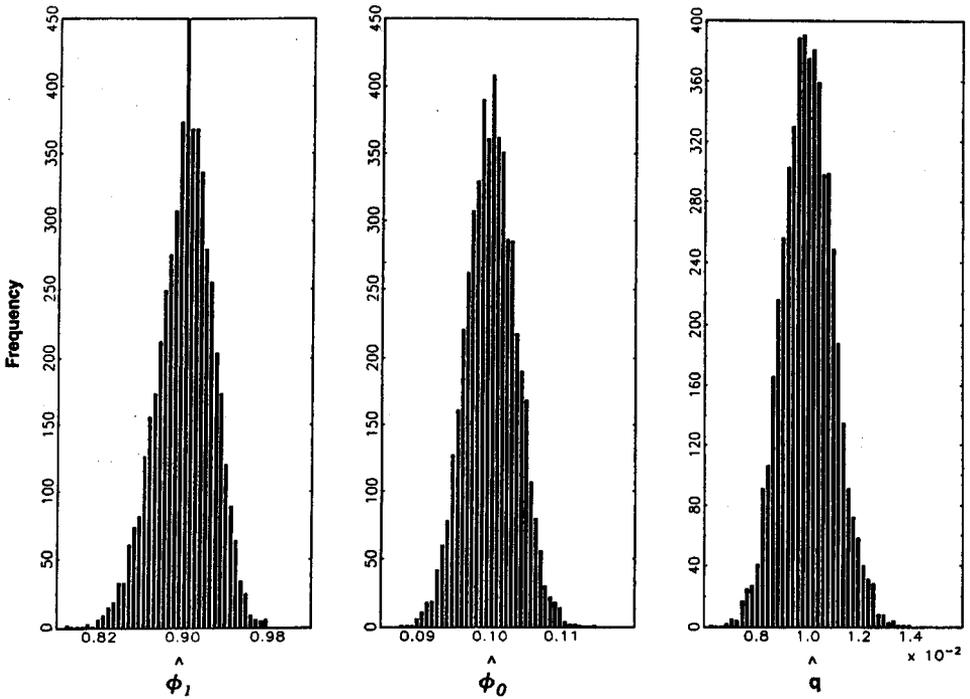


Fig. 1. Histograms of $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} for samples with 5893 families, 10% of which have case probands. The corresponding parameter values are $\phi_1 = 0.9, \phi_0 = 0.1$, and $q = 0.01$.

With higher allele frequencies ($q=0.10$), smaller samples are needed to achieve good precision. Indeed, only 577 families are required with $\phi_1 = 0.9, \phi_0 = 0.1$ and $q = 0.1$. As expected, there is no evidence of bias and confidence intervals have near-nominal coverage in cases with $\pm 5\%$ precision (simulations 4 and 5 in Table 1), and

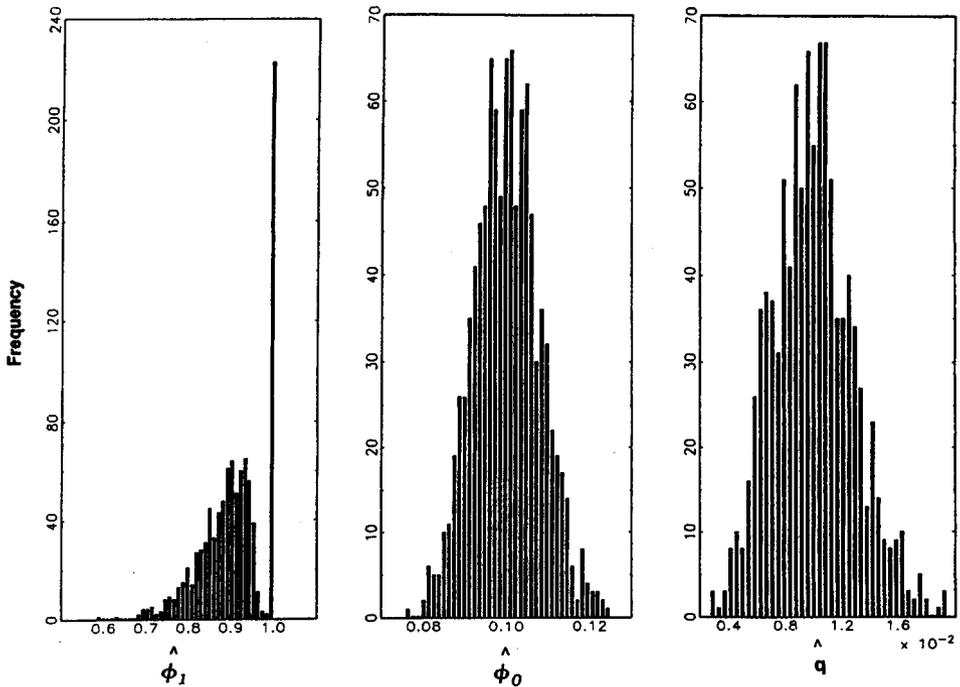


Fig. 2. Histograms of $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} for samples with 737 families, 10% of which have case probands. The corresponding parameter values are $\phi_1 = 0.9$, $\phi_0 = 0.1$, and $q = 0.01$.

corresponding histograms indicate symmetric distributions of $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} in these cases (not shown).

3.2. Selection bias

A major bias can be introduced in kin-cohort analyses if persons with diseased relatives are more likely to volunteer to be probands than persons without diseased relatives (Struwing et al., 1997; Wacholder et al., 1998; GPBC). To illustrate this bias, consider an example with $\phi_1 = 0.90$, $\phi_0 = 0.10$ and $q = 0.1$. Suppose that mothers with an affected daughter are twice as likely to volunteer to be probands as are mothers with no affected daughters and that mothers with two affected daughters are 4 times as likely to volunteer, regardless of whether the mother is affected or not. A simulation with 5000 repetitions and 1154 families, of which 10% have affected probands, yielded average estimates (\pm standard errors) $0.944(\pm 0.009)$ for ϕ_1 , $0.150(\pm 0.015)$ for ϕ_0 and $0.205(\pm 0.015)$ for q . Thus, the penetrances ϕ_1 and ϕ_0 are substantially overestimated, as is the allele frequency q . These biases reflect the increased burden of disease in the sampled data that results from biased selection of participating probands.

Table 2

Effects of misclassifying phenotypes of relatives with true $\phi_1 = 0.9$, $\phi_0 = 0.1$ and $q = 0.01^a$

Sensitivity	Specificity	Mean estimates			Coverage of confidence intervals for ϕ_1	
		$\hat{\phi}_1$	$\hat{\phi}_0$	\hat{q}	Wald	Likelihood ratio
		1.00	1.00	0.900		
0.90	1.00	0.873	0.090	0.009	0.884	0.846
1.00	0.90	0.945	0.184	0.017	0.230	0.301
0.90	0.90	0.933	0.174	0.016	0.534	0.617

^aEach experiment was based on 1000 simulated repetitions. Each simulation included 5893 families, with 10% having diseased probands; these sample sizes were designed to achieve $\pm 5\%$ precision on $\hat{\phi}_1$.

3.3. Misclassification of disease status of relatives of probands

An advantage of the kin-cohort design is that one can sometimes obtain information on the disease status (or phenotypes) of relatives simply by interviewing the proband. This approach can lead to misclassification of relatives' phenotypes, however. To investigate such misclassification, we define sensitivity as the probability that a proband will correctly report a diseased relative as diseased and specificity as the probability that a proband will correctly report a non-diseased relative as non-diseased. We consider the effects of imperfect sensitivity and specificity for $\phi_1 = 0.9$, $\phi_0 = 0.1$, $q = 0.01$ and with 5893 families selected to yield precision $\pm 5\%$ for ϕ_1 . Table 2 reports the results of 1000 simulations for each combination of sensitivity and specificity.

With sensitivity 0.9 and specificity 1.0, estimates of ϕ_0 and q_0 are downwardly biased by 10% (Table 2), and estimates of ϕ_1 by 3%. With sensitivity 1.0 and specificity 0.9, the average values of $\hat{\phi}_1$, $\hat{\phi}_0$ and \hat{q} are 0.945, 0.184, and 0.017, indicating a substantial upward bias in each parameter estimate. A similar upward bias is seen when both the sensitivity and specificity are 0.9 (Table 2). In this setting

$$P(Y_0 = 1) = \phi_1 P(g_0 = 1) + \phi_0 P(g_0 = 0) = 0.9(1 - 0.99^2) + (0.1)(0.99^2) = 0.1159.$$

Thus even a modest decrease in specificity leads to an important number of false positive results and inflated estimates of penetrance and allele frequency, q .

3.4. Residual familial correlations induced by factors other than the mutation under study

The analyses of kin-cohort data in GPBC, Struewing et al. (1997) and Wacholder et al. (1998) explicitly or implicitly assume conditional independence of phenotypes, given genotypes (Assumption (A3) in Section 1). This assumption is also used as the starting point for most segregation analyses, although some models for segregation analysis also allow for the possibility of residual familial aggregation from other sources (see e.g. Li and Thompson, 1997). In unpublished work, by R. Carroll, D. Pee,

Table 3

Effects of residual familial correlations induced by factors other than the mutation (A) under study, with marginal penetrances $P(Y = 1|g = 1) = 0.9$ and $P(Y = 1|g = 0) = 0.1$ and with $q = 0.01^a$

τ^2	$\{1 + \exp(-\mu_1)\}^{-1}$	$\{1 + \exp(-\mu_0)\}^{-1}$	Average value of			Coverage of confidence interval for ϕ_1	
			$\hat{\phi}_1$	$\hat{\phi}_0$	\hat{q}	Wald	Likelihood ratio
0	0.900	0.100	0.900	0.100	0.010	0.932	0.945
0.5	0.916	0.084	0.910	0.098	0.011	0.941	0.938
1	0.928	0.072	0.917	0.096	0.011	0.928	0.917
2	0.947	0.053	0.933	0.093	0.012	0.933	0.889
4	0.968	0.032	0.948	0.087	0.015	0.909	0.818

^aEach experiment included 1000 simulations with probability of disease determined by the logistic random effects model in Section 3.3 with parameters μ_1 and τ^2 for $g = 1$ and μ_0 and τ^2 for $g = 0$.

J. Benichou and M.H. Gail, it was shown that there is evidence of such residual correlation in a sample of the data analyzed by Struewing et al. (1997). To the extent that family members share exposure to measured risk factors for the disease in question, residual correlation can be accounted for by covariate adjustment using regression models and by assuming that phenotypes are conditionally independent given genotypes and measured covariates. Residual familial correlation can also result from unmeasured shared exposures or risk factors, however, such as other unidentified mutations segregating in the family or shared but unmeasured dietary habits.

To allow for such unmeasured familial factors, we considered the logistic random effects model

$$P(Y = 1|g, b) = \{1 + \exp(-\mu_g + b)\}^{-1}, \tag{2}$$

where μ_1 and μ_0 correspond to $g = 1$ or 0 and where b is a mean zero normal variate with variance τ^2 . Values of b are drawn independently for each family. Each family member is influenced by the same random familial effect, b , in addition to his or her genotype. In particular, Y_{11} , Y_{12} , and Y_0 are each influenced by the random familial effect. For each value of τ^2 , we calculated μ_1 and μ_0 so that the marginal probabilities $P(Y = 1|g = 1) = 0.9$ and $P(Y = 1|g = 0) = 0.1$, and we assumed $q = 0.01$. Transformed values of μ_1 and μ_0 are shown in Table 3 and indicate that as τ^2 increases, more extreme values of μ_1 and μ_0 are needed to maintain the marginal probabilities above. For $g_{11} = 1$, $g_{12} = 1$, the intra-familial correlations comparing two family members are 0.047, 0.092, 0.171, and 0.287 for $\tau^2 = 0.5, 1, 2$ and 4 , respectively. For $g_{11} = 1$, $g_{12} = 0$ these correlations are 0.036, 0.058, 0.083, and 0.102, and for $g_{11} = 0$, $g_{12} = 0$, they are 0.047, 0.092, 0.171, and 0.287.

As the value of τ^2 increases, estimates of ϕ_1 , ϕ_0 and q based on conditional independence become increasingly more biased (Table 3). In fact, $\hat{\phi}_1$ overestimates ϕ_1 , \hat{q}_1 overestimates q , and $\hat{\phi}_0$ underestimates ϕ_0 . Because the simple genetic model is forced to account for genetic and residual familial correlation, it exaggerates the genetic effects by overestimating $P(Y|g = 1)$, underestimating $P(Y|g = 0)$ and overestimating the allele frequency, q .

We performed a similar simulation to study the effect of an unidentified gene with alleles c and C segregating independently of the gene with alleles a and A . We assumed $P(A)=0.01=q$, as before and $P(C)=0.01626$. We let $h=1$ if the subject had genotype CC or Cc , and 0 if cc . To maintain the original probabilities $P(Y=1|g=1)=0.9$ and $P(Y=1|g=0)=0.1$, we assumed that the joint effects of the two genes satisfied $P(Y=1|g=0, h=0)=0.080$, $P(Y=1|g=0, h=1)=0.700$, $P(Y=1|g=1, h=0)=0.897$, and $P(Y=1|g=1, h=1)=0.990$. The average values of $\hat{\phi}_1$, $\hat{\phi}_0$ and \hat{q} (with standard errors) in 5000 simulations under the naive model ignoring the second gene were 0.913(0.023), 0.098(0.003), and 0.0108(0.0011). Thus, the bias is very slight (and not statistically significant in these studies), because the ignored allele, C , is so rare. This result is different from that for a conventional segregation analysis, in which the combined effects of A and C and a combined allele frequency will be estimated, rather than the average effect and allele frequency of A alone.

We also considered a common mutation with $P(C)=0.5$ and with $P(Y=1|g=0, h=0)=0.05714$, $P(Y=1|g=0, h=1)=0.11429$, $P(Y=1|g=1, h=0)=0.62857$, and $P(Y=1|g=1, h=1)=0.99048$. These parameters yield marginal probabilities $P(Y=1|g=1)=0.9$ and $P(Y=1|g=0)=0.1$, as before. Simulations based on 5000 repetitions gave average estimates of $\hat{\phi}_1$, $\hat{\phi}_0$ and \hat{q} (with standard errors) of 0.900(0.027), 0.100(0.003), and 0.010(0.001). Thus, even ignoring a common gene in this example induced negligible bias in estimates of $P(Y=1|g=1)$, $P(Y=1|g=0)$, and q .

3.5. Violation of Hardy–Weinberg equilibrium because of stratification

We examined the bias that results when the population is divided into two strata, within each of which mating is at random, but between which no mating occurs. In such a population, Hardy–Weinberg equilibrium holds only within strata. We can calculate the asymptotic bias that arises from assuming Hardy–Weinberg equilibrium in the entire population by taking expectations of the misspecified score equations from Eq. (1) with respect to the true stratified probability distribution.

To study the case of a rare allele, we assumed stratum frequencies 0.2 and 0.8, with respective allele frequencies 0.046299 and 0.001131 chosen to preserve $P(g=1)=1-(1-0.01)^2=0.01990$ in the entire population. The allele frequency in the whole population is 0.01016. The mles $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} converge, respectively, to 0.9025, 0.0997, and 0.0998, which are quite close to the true values 0.9, 0.1 and 0.01016. Thus, for rare alleles, such stratification and failure of Hardy–Weinberg equilibrium induce little bias.

With larger allele frequencies, however, the bias is more noticeable. If the respective allele frequencies in the previous strata were 0.630726 and 0.010854, then the allele frequency in the whole population would be $q=0.1348$. In this case $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} converge, respectively, to 0.914, 0.118, and 0.1125 instead of the proper values, 0.9, 0.1, and 0.1348. Thus stratification has noticeable but small effect on $\hat{\phi}_1$, $\hat{\phi}_0$, and \hat{q} , even for larger allele frequencies.

4. Discussion

In this paper we have examined violations of some of the assumptions underlying maximum likelihood analysis of the kin-cohort design. Although these numerical studies cover a limited range of the parameter space and concentrate on the important case of rare alleles with high penetrance, they indicate that some violations seriously distort the findings. In particular, if probands tend to self-select for study because they have diseased relatives, estimates of penetrance and allele frequencies can be seriously upwardly biased. Population-based case-control studies and cohort studies are much less affected by such selection bias, as discussed by GPBC. Misclassification of relatives' phenotypes can also induce serious bias, especially when unaffected relatives are reported as diseased (Table 2).

Other violations are less serious, such as the effects of failure of Hardy-Weinberg equilibrium from stratification of the population (Section 3.5).

Large samples may be required to attain good precision for estimates of $\hat{\phi}_1$, but $\hat{\phi}_1$, $\hat{\phi}_0$ and \hat{q} seem to be nearly unbiased estimators even for smaller samples for which $\hat{\phi}_1$ is less precise and is not normally distributed. If samples are small enough that $\hat{\phi}_1$ falls on the boundary $\hat{\phi}_1 = 1$ frequently, then the coverage of the Wald confidence interval can drop below nominal levels. A test-based confidence interval derived from the profile likelihood ratio test for ϕ_1 had near-nominal coverage in such cases, however (Table 1).

Failure of the conditional independence Assumption (A3) in Section 1, can lead to bias. Strong familial random effects induce overestimates of $P(Y = 1|g = 1)$ and q and underestimates of $P(Y = 1|g = 0)$ (Table 3). If such effects are induced by rare co-segregating genes, they have little influence on estimates of $P(Y = 1|g = 1)$, $P(Y = 1|g = 0)$ and q . In this respect, a kin-cohort design with measurements of the probands' genotypes, AA , Aa or aa , is more robust than a simple segregation analysis, which will lump the effects of unidentified co-segregating genes together with the effects of the gene of interest. It is encouraging that ignoring a more common co-segregating gene with allele frequency 0.5 did not induce appreciable bias in estimates of $P(Y = 1|g = 1)$, $P(Y = 1|g = 0)$, and q . Nevertheless, the possibility of residual familial correlation cannot be ignored (Table 3), and more elaborate models that allow for residual familial correlation beyond the gene of interest, such as the frailty model of Li and Thompson (1997), could be incorporated in kin-cohort analyses to weaken the conditional independence assumption.

Acknowledgements

This research was supported by a grant from National Cancer Institute (CA-57030), and by the Texas A&M Center for Environmental and Rural Health via a grant from the National Institute of Environmental Health Sciences (P30-ES09106).

References

- Claus, E.B., Risch, N., Thompson, W.D., 1991. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am. J. Hum. Genet.* 48, 232–242.
- Cornfield, J., 1951. A method for estimating comparative rates from clinical data: applications to cancer of the lung, breast and cervix. *J. Natl. Cancer Inst.* 11, 1269–1275.
- Elston, R.C., Stewart, J., 1971. A general model for the genetic analysis of pedigree data. *Hum. Hered.* 21, 523–542.
- Gail, M.H., Pee, D., Benichou, J., Carroll, R., 1999. Designing studies to estimate the penetrance of an identified autosomal dominant mutation: cohort, case-control and genotyped-proband designs. *Genet. Epidemiol.* 16, 15–39.
- GAUSS, Version 3.1, 1993. Systems and Graphics Manual. Aptec Systems Inc., Maple Valley, Washington, DC.
- Li, H., Thompson, E., 1997. Semiparametric estimation of a major gene and family-specific random effects for age of onset. *Biometrics* 53, 282–293.
- Patefield, W.M., 1977. On the maximized likelihood function. *Indian J. Statist.* 39 (Series B), 92–96.
- Struewing, J.P., Hartge, P., Wacholder, S., Baker, S.M., Berlin, M., McAdams, M., Tucker, M.A., 1997. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *N. Engl. J. Med.* 336, 1401–1408.
- Wacholder, S., Hartge, P., Struewing, J.P., Pee, D., McAdams, M., Brody, L., Tucker, M., 1998. The kin-cohort study for estimating penetrance. *Am. J. Epidemiol.* 148, 623–630.