Published in final edited form as:

Genet Epidemiol. 1995 ; 12(6): 859–862. doi:10.1002/gepi.1370120655.

Variability of Genotype-specific Penetrance Probabilities in the Calculation of Risk Support Intervals

Suzanne M. Leal and

Department of Otolaryngology, University of Tübingen, Germany;

Jurg Ott

Department of Psychiatry, Columbia University, New York

Abstract

Previously, a maximum likelihood method was described to construct a support interval for the risk. This method is extended to incorporate genotype specific penetrance probabilities in the calculation of a risk support interval. As an empirical example, the support interval for the risk is calculated for a member of a published breast-ovarian cancer kindred.

Keywords

genetic counseling; support interval; genetic risk; phenotype risk

INTRODUCTION

In a Mendelian trait, the genetic risk is the conditional probability that an individual has the genetic susceptible genotype given both phenotype and genotype information for all available pedigree members. Genetic risks may be based on the pedigree likelihood [Elston and Stewart, 1971]. In addition to such genotype risks, a phenotype risk may be defined as the conditional probability of developing the trait. With incomplete penetrance and absence of phenocopies, the phenotype risk is smaller than the corresponding genotype risk. Generally, however, phenocopies as well as genetic cases contribute to the phenotype risk.

The precision of risk estimates is dependent on the accuracy of the parameters used in their evaluation. Usually risks are computed under the assumption that genetic parameters are known without error. Uncertainty in the accuracy of parameter estimates renders uncertainty in the risk. Therefore, to evaluate the accuracy of a risk it is critical to calculate either a confidence or support interval for the risk, [Smith, 1971; Lange, 1986; Rogatko, 1988; Weeks and Ott, 1989; Suthers and Wilson, 1990; Ott, 1991; Leal and Ott, 1994].

Leal and Ott

Previously, we described a method to construct support intervals (SIs) for genetic risks working in a maximum likelihood framework [Leal and Ott, 1994]. Briefly, the method

allows for parameters to vary in their support intervals. For each combination of parameter values so obtained, a risk is calculated whose associated log likelihood is equal to the log likelihood at the given parameter values. All those risk values with a log likelihood within m units of the maximum log likelihood form the risk support interval. Here, this method is expanded to allow for variability of genotype-specific penetrances when the age at disease onset is normally distributed. As an empirical example, the SIs for the phenotype and genotype risk will be calculated below for a member of a breast-ovarian cancer kindred using two markers (D17S250 and D17S588) which are linked to the BRCA 1 locus.

METHODS

In genetic counseling situations, one generally works with a single pedigree. Usually, parameter estimates must be obtained from previously published results. A maximum likelihood method to construct an SI for the risk under these circumstances was previously described [Leal and Ott, 1994]. In principle, we rely on published support intervals. If these are unavailable, we construct them by one of several methods using information in published reports.

Below genotype-specific penetrance probabilities are incorporated in the calculation of SIs for the genotype and phenotype risk. Approximate m -unit SIs are constructed around the mean age of disease onset, μ , and lifetime penetrances, λ , each for disease gene carriers and noncarriers. The calculation of maximum and joint log likelihoods for all parameters is carried out as previously described [Leal and Ott, 1994], except that here, the estimates, μ , for age at disease onset are taken to follow a normal distribution while all other parameter estimates are binomially distributed.

The penetrance probabilities are the genotype-specific cumulative risk for unaffected and affected individuals when age of onset is unknown, and genotype-specific density for affected individuals when age of onset is known.

At this point, each parameter is varied within its SI. When the joint log likelihood for a set of parameter values falls within m units of maximum log likelihood, the genotype-specific penetrance probabilities are calculated for each liability class and the risk is calculated with the aid of MLINK [Lathrop and Lalouel, 1984]. The phenotype risk is also computed using a specific cumulative penetrance liability class. The highest and lowest (genotype and phenotype) risks so obtained are taken to be the endpoints of the (genotype and phenotype) risk SI.

EXAMPLE

As an empirical example, 2-unit SIs for the phenotype and genotype risk were calculated for individual 405, an unaffected 52-year-old female who is a member of the breast-ovarian cancer kindred CRCIOI [Smith et al., 1993], given her current age.

The following parameter values were used: male map distance between the two flanking markers D17S250 and S17S588 = 9.8 cM Haldane; female-to-male map distance ratio = 2.0; the disease locus map position, x = 4.5 cM centromeric to D17S588, the SI (maximum lod

score- 1) for this map position ranges 2.9 cM proximal from D17S588 to 3.6 cM distal to D17S250 (length of SI 3.28 cM); the proportion of linked families, $\alpha = 1.0$ with an SI of (0.71,1.0) [Easton et al., 1993; Bishop, personal communication]; the age of onset was assumed to be normally distributed with a mean age of onset for breast cancer gene carriers, $\mu_{AA} = \mu_{Aa} = 55.435$ (standard error = 1.742) and noncarriers, $\mu_{aa} = 68.990$ (standard error = 1.532), the standard deviation, $\sigma_{AA} = \sigma_{Aa} = \sigma_{aa} = 15.387$; the lifetime penetrance for breast cancer gene carriers, $\lambda_{AA} = \sigma_{Aa} = 0.928$ (standard error = 0.163) and noncarriers, $\lambda_{aa} = 0.1$ (standard error = 0.009); and the disease gene frequency, $q = 0.0033$ (standard error = 0.0012) [Claus et al., 1991].

As previously described by Easton et al. [1993], affected and unaffected individuals were assigned to disease and age specific liability classes with all males and ovarian cancer cases assigned to the youngest age class. A total of 14 age specific liability classes were formed, 7 for unaffected and 7 for affected individuals (ages < 30, 30–39, 40–49, 50–59, 60–69, 70–79, and >80) using the middle of each age class (25, 35, etc).

The parameters, x , α , q , μ_{AA} , μ_{aa} , λ_{AA} , and λ_{aa} were each varied while all other parameters were held fixed. Whenever the joint log likelihood fell within 2 units of the maximum log likelihood, the risk for the counselee was calculated under heterogeneity as a weighted average of the risk under linkage and no linkage [Weeks and Ott, 1989]. The phenotype risk (that individual 405 will be affected within her lifetime [< 80 years of age]) was calculated using the cumulative penetrance age class 70–79.

The highest and lowest (genotype and phenotype) risks were taken to be the upper and lower bounds for the (genotype and phenotype) SI. Point (phenotype and genotype) risk estimates were calculated under homogeneity, $\alpha = 1.0$ using the point estimates for all parameters.

RESULTS

The SI for the genotype risk that individual 405 carries the BRCA1 susceptibility allele is (0.0%, 14.5%) and the SI for the phenotype risk is (5.9%, 19.4%). The point estimates for the genotype and phenotype risks are 0.021 and 0.084, respectively.

DISCUSSION

The calculation of SIs enables genetic counselors to determine the reliability of risk estimates. An SI for the risk can help to determine the accuracy of the risk estimate, where a wide SI reflects an inaccurate point estimate.

In the method described, only “statistical” or sampling variability of the estimate is allowed for in the calculation of SIs for the risk. Phenotypic errors, marker map position errors, and systematic bias will also affect the accuracy of the risk; however, these errors are not taken into account in this method.

For this example, the SI for the genetic risk is wide (0.0%, 14.5%). However, the upper bound of the SI for the risk is lower than the counselee's genotype risk ($R = 40.5\%$) if no marker data were available. The SI for the phenotype risk (5.9%, 19.4%) reflects the

variability of the genotype-specific cumulative penetrances for breast cancer gene and nongene carriers. It should be noted that often SIs may be wider than one would expect, especially under linkage heterogeneity [Leal and Ott, 1994].

The computer program, RISKSI, which implements the procedures outlined above, and an auxiliary program, RISKPREP, which aids users in the creation of the necessary data file are available on the anonymous ftp site linkage.cpmc.columbia.edu.

ACKNOWLEDGMENTS

This material is based upon work supported by US Army Medical Research Acquisition Activity under award #DAMD17-94-J-4406. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the US Army Medical Research Acquisition Activity. We also acknowledge support by the Deutsche Forschungsgemeinschaft (Klifo Horforschung; Zel49/6-1).

REFERENCES

- Claus E, Risch N, Thompson WD (1991): Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 48:232-242. [PubMed: 1990835]
- Easton DF, Bishop DT, Ford D, Crockford GP and the Breast Cancer Linkage Consortium (1993): Genetic linkage analysis in familial breast and ovarian cancer: Results from 214 families. *Am J Hum Genet* 52:678-701. [PubMed: 8460634]
- Elston RC, Stewart J (1971): A general model for the analysis of pedigree data. *Hum Hered* 21:523-442.
- Lange K (1986): Approximate Confidence intervals for risk prediction in genetic counseling. *Am J Hum Genet* 38:681-687. [PubMed: 3459355]
- Lathrop GM, Lalouel JM (1984): Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet* 36:460-465. [PubMed: 6585139]
- Leal SM, Ott J (1994): A likelihood approach to calculating risk support intervals. *Am J Hum Genet* 54:913-917. [PubMed: 8178830]
- Ott J (1991): "Analysis of Human Genetic Linkage" Baltimore: Johns Hopkins University Press.
- Rogatko A (1988): Evaluating the uncertainty of risk predication in genetic counseling: A Bayesian approach. *Am J Med Genet* 31:513-519. [PubMed: 3228134]
- Smith C (1971): Recurrence risk for multifactorial inheritance. *Am J Hum Genet* 23:578-588. [PubMed: 5132066]
- Smith SA, Easton DF, Ford D, Peto J, Anderson K, Averill M, Stratton M, Ponder M, Pye C, Ponder BJA (1993): Genetic heterogeneity and localization of a familial breast-ovarian cancer gene on chromosome 17q12-q21. *Am J Hum Genet* 52:767-776. [PubMed: 8460642]
- Suthers GK, Wilson SR (1990): Genetic counseling in rare syndromes: a resampling method for determining an approximate confidence interval for gene location with linkage data from a single pedigree. *Am J Hum Genet* 47:53-61. [PubMed: 2349950]
- Weeks DE, Ott J (1989): Risk calculations under heterogeneity. *Am J Hum Genet* 45:819-821. [PubMed: 2816945]