AD_____

GRANT NO: DAMD17-94-J-4406

TITLE: Statistical Genetic Methods for Localizing Multiple Breast Cancer Genes

PRINCIPAL INVESTIGATOR(S): Doctor Jurg Ott

CONTRACTING ORGANIZATION: Columbia University
in the City of New York
New York, New York 10032

REPORT DATE: September 1995

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1995	Annual 1 Sep 94 - 31 Aug 95

4. TITLE AND SUBTITLE
Statistical Genetic Methods for Localizing Multiple Breast Cancer Genes

5. FUNDING NUMBERS
DAMD17-94-J-4406

6. AUTHOR(S)
Doctor Jurg Ott

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Columbia University
New York, New York 10032

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Genetic risks are often computed as a single number (a so-called point estimate) under the assumption that all parameters involved in the calculation of the risk are known without error. In previous work, a method was developed to allow for variability in these parameters. Technically, by the use of the maximum likelihood method, a support interval (approximate confidence interval) is constructed for the risk. This method has been extended to incorporate age-dependent penetrances in the calculation of risk support intervals. As an empirical example, the support interval for the risk is calculated for a member of a published breast-ovarian cancer kindred.

14. SUBJECT TERMS
breast cancer
Risk, genotype, phenotype, support interval, confidence interval, likelihood

15. NUMBER OF PAGES
7

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet *optical scanning requirements*.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- **DOD** - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank.
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Jurgon 9/22/95

PI - Signature Date

(none of the above apply to this project)

TABLE OF CONTENTS

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 as originally proposed by 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, in order to evaluate the accuracy of a risk it is critical to calculate either a confidence or support interval for the risk (Weeks and Ott 1989; Ott, 1991).

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. Under this grant, as proposed, 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 have been calculated for a member of a breast-ovarian cancer kindred using two markers (D17S250 and D17S588) which are linked to the BRCA1 locus.

BODY

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.

As proposed under this grant, genotype-specific penetrance probabilities are incorporated in the calculation of SIs for the genotype and phenotype risk in the following manner: 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. In the likelihood calculations, 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. These changes have been implemented in the RISKSI program and described in a manuscript (Leal and Ott 1995).

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 CRC101 (Smith et al. 1993), given her current age. Technical details may be found in the manuscript (Leal and Ott 1995). The resulting SI for the genotype risk that she carries the BRCA1 susceptibility allele ranges from 0 through 14.5%, and the SI for her phenotype risk is between 5.9% and 19.4%. The point estimates for the genotype and phenotype risk are 2.1% and 8.4%, respectively.

CONCLUSIONS

The calculation of support intervals 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.

REFERENCES

Elston RC, Stewart J (1971): A general model for the analysis of pedigree data. Hum Hered 21:523-542

Lathrop GM, Lalouel JM (1984): Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36:460-465

Leal SM, Ott J (1994): A likelihood approach to calculating risk support intervals. Am J Hum Genet 54:913-917

Leal SM, Ott J (1995) Variability of Genotype Specific Penetrance Probabilities in the Calculation of Risk Support Intervals. Genet Epidemiol (in press)

Ott J (1991) "Analysis of human genetic linkage." Baltimore: Johns Hopkins University Press

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

Weeks DE, Ott J (1989) Risk calculations under heterogeneity. Am J Hum Genet 45:819-821

APPENDIX

N/A