Genetic Epidemiology 12:317–336

Abstracts From the Fourth Annual Meeting of the International Genetic Epidemiology Society

Snowbird Resort
Salt Lake City, Utah
June 20–22, 1995

Hosted by the University of Washington and the University of Utah Conferences and Institutes

1

Two-locus maximum lod score analysis of a polygenic trait: evidence for locus heterogeneity at IDDM1 and IDDM4 in type 1 diabetes

Heather J. Cordell, John A. Todd, Simon T. Bennett, Yoshihiko Kawaguchi and Martin Farrall
The Wellcome Trust Centre for Human Genetics, Windmill Road, Oxford OX3 7BN, UK

To investigate the genetic component of multifactorial diseases, models involving the joint action of several disease loci are important. These models give increased power to detect an effect and a greater understanding of aetiological mechanisms. Here we present an extension of the maximum lod score (MLS) method of Risch (1990c) which allows the simultaneous detection and modelling of two unlinked disease loci. Genetic constraints on the identical by descent (ibd) sharing probabilities are derived and the size and power of the test statistics is investigated. The method is applied to affected sib-pair data and the joint effects of *IDDM1* (HLA) and *IDDM2* (the *INS* VNTR), and *IDDM1* and *IDDM4* (*FGF3*), are assesed with relation to type 1 diabetes (IDDM). In the presence of genetic heterogeneity there is seen to be a significant advantage in analysing more than one locus simultaneously. The results indicate that the effects at *IDDM1* and *IDDM2* are well described by a multiplicative genetic model, while those at *IDDM1* and *IDDM4* follow a heterogeneity model. These data suggest that there are two forms of IDDM, one in which *IDDM1* is predominant, and another in which the action of *IDDM4* is important.

2

Complex segregation analysis of nasopharyngeal carcinoma (NPC) in Taiwan. Y.F. Wang[1,2], T.H. Beaty[2], E.L. Harris[3], M.M. Hsu[4], S.R. Diehl[1], C.J. Chen[4]. [1]National Institute of Dental Research, NIH, Bethesda, Maryland. [2]The Johns Hopkins University, Baltimore, Maryland. [3]Kaiser Permanente Center for Health Research, Portland, Oregon. [4]National Taiwan University, Taipei, Taiwan, R.O.C.

A family study of NPC was carried out in Taiwan to define the nature of the familial aggregation of NPC, and determine whether or not segregation of a major gene could explain the distribution of NPC in families. Detailed information about family data, smoking habits, and consumption of salted fish during childhood was collected from 750 NPC probands. Segregation analysis of a truncated trait with a logistic probability density function for age of onset was performed

using the S.A.G.E. package (REGTL). A series of genetic and non-genetic models was examined. Results suggested that familial clustering of NPC can be best explained by a Mendelian recessive locus with a gender specific susceptibility, where consumption of salted fish during childhood and the number of unaffected preceding sibs also affects risk. Under this model, the frequency of the high risk allele was 0.18, and the lifetime susceptibility of NPC was 0.091 for females and 0.178 for males. Penetrance of this putative high risk allele is modest at best: by age 80, only 8.1% of females and 15.7% of males with the high risk genotype would develop NPC (in the absence of other risk factors), compared to 0.04% for females and 0.07% for males with low risk genotypes. Persons who consumed salted fish during childhood have a slightly higher penetrance (8.8% for females, and 17.2% for males by age 80 for the high risk genotype). Although adding the number of preceding sibs improved the fit of the model significantly, the increased risk associated with a prior non-affected sib was small. There was no evidence of a statistically significant effect of smoking. The best-fitting model of inheritance for age-of-onset of NPC further suggested that the high risk gene accounted for over 90% of NPC in this population.

3

REGRESSIVE LOGISTIC MODELLING OF FAMILIAL AGGREGATION FOR ASTHMA IN 7,394 POPULATION BASED NUCLEAR FAMILIES

Mark A. Jenkins, John L. Hopper
The University of Melbourne,Department of Public Health & Community Medicine,
200 Berkeley Street, Carlton, Vic. 3053, Aust.

The aim of this population-based study was to determine whether asthma aggregates in families and if so, whether this aggregation was consistent with environmental and/or genetic factors. Data was from 7,394 two-generation families containing 41,506 individuals from the 1968 Tasmanian Asthma Survey, in which all Tasmanian school children born in 1961 (probands) were surveyed by a respiratory questionnaire completed by their parents. Similar data was obtained for the parents and all siblings of the probands.For a child, having ever had asthma was predicted by a parent having ever had asthma,with an odds ratio (OR) of 3.60(95%confidence interval 3.14-4.13)for mother and 3.22(2.81-3.70) for father, and by a sibling having ever had asthma Regressive logistic modelling of asthma prevalence showed that the data was consistent with the existence of an unmeasured factor shared by siblings,evident in 15±2% of families and associated with an odds ratio of 9.69(8.27-11.32). Segregation analysis predicted that a codominant model with a population disease allele frequency of 19±0.3% was more consistent with the data than a dominant or recessive model.On the other hand,under a dominant locus model there was evidence for residual parent-offspring and sibling effects.We conclude that there may be a number of genetic loci influencing asthma susceptibility, and that dominant transmission is the more likely mode of transmission for at least a majority of these loci.

4

The *IGF1* locus is a major determinant of serum osteocalcin levels in Mexican Americans.

J. Blangero, P.B. Samollow, M.B. Rocha, J.E. Hixson and J. Rogers, Southwest Foundation for Biomedical Research, San Antonio, TX 78228.

Serum concentration of osteocalcin (sOC) is an important marker of bone formation and correlates with osteoporosis risk. However, little is known about the genetic determinants of sOC level in humans. Recent studies suggest that genetic variation at the vitamin D receptor (*VDR*) locus may be responsible for some of the observed quantitative variation in osteocalcin level. To assess the role of genetic factors in the determination of this trait, we measured sOC levels in 481 Mexican Americans who were members of 26 randomly ascertained pedigrees. Polymorphic genetic markers at two candidate loci, *VDR* and the insulin-like growth factor I (*IGF1*) locus, were also typed.

Quantitative trait linkage analysis was performed to test whether the candidate loci influenced sOC level. A general variance component method allowing for oligogenic determination of the trait was used to test for linkage. Marker-specific matrices containing the pairwise probabilities of identity-by-descent were estimated using a Monte Carlo algorithm and used to model the variances due to genes linked to the marker loci.

The statistical genetic analyses revealed no evidence for genes influencing sOC at or near the *VDR* locus ($p = 0.365$). However, a significant linkage to the *IGF1* locus was found ($p = 0.018$). Variance component estimates showed that the *IGF1* gene (or a nearby gene) accounts for 35% of the phenotypic variation in sOC levels in this population. This linkage to the *IGF1* locus accounts for all of the observed genetic variance in sOC.

Supported by NIH grants DK42273 and HL45522.

5

Apolipoprotein E, family history of dementia, sex, and age at onset of Alzheimer's disease.

SC Waring, WA Rocca, DJ Schaid, RC Petersen, EG Tangalos, SN Thibodeau, E Kokmen
Departments of Health Sciences Research, Neurology, Community Medicine, and Laboratory Medicine and Pathology, Mayo Clinic and Foundation, Rochester, MN

We evaluated the effects of apolipoprotein E (apoE), family history of dementia (dementia in at least one first-degree relative), and sex on age at onset of Alzheimer's disease (AD). All patients with a clinical diagnosis of AD and apoE genotyping (n=142) were identified from the Mayo Clinic Alzheimer's Disease Patient Registry, which is a community-based prospective series (mainly Caucasians). Kaplan-Meier curves for all patients showed a significant pattern of decreasing age at onset of AD by increasing number of apoE ε4 alleles (p=0.0002). The table shows the distributions of median age at onset by genotype and family history of dementia (number of patients in parentheses):

	Total	FamHx +	FamHx -
ε4/ε4	74.1 (13)	69.1 (9)	75.1 (4)
ε4/ε2, 3	78.1 (59)	77.8 (34)	80.6 (25)
ε2/ε2, 3	84.6 (7)	82.7 (3)	85.6 (4)
ε3/ε3	82.6 (63)	81.9 (23)	83.3 (40)

These effects did not vary by sex across all strata. We further investigated all univariate and multivariate main effects and first order interactions of apoE, sex, and family history using Cox proportional hazards models. The final model indicated that family history (hazard ratio 1.50, p=0.0217) and number of ε4 alleles (hazard ratios: 2 alleles=2.69; 1 allele=1.64; p=0.0003) predict earlier age at onset of AD. By contrast, the ε2 allele does not show the significant delaying effect reported by others. We found no statistically significant interaction between apoE and family history. Our findings indicate that apoE ε4 and family history of dementia have an independent effect on age at onset of AD in this population.

6

Bias of the association between Apolipoprotein E (APOE) and Alzheimer's disease by differential survival related to the APOE genotype. CM an Duijn,[1] P De Knijff,[2] A Wehnert,[3] J De Voecht,[3] JB Bronzova,[1] LM Havekes,[2] A Hofman,[1] C Van Broeckhoven.[3] [1]Dep of Epidemiology & Biostatistics, Erasmus University Medical School, Rotterdam, The Netherlands; [2]TNO Institute of Prevention and Health, Gaubius Laboratory, Leiden, The Netherlands; [3]Neurogenetics Laboratory, Born Bunge Foundation, University of Antwerp, Antwerp, Belgium

It was suggested that in contrast to the E4 allele, the E2 allele of the Apolipoprotein E gene (APOE*2) has a protective effect for late-onset Alzheimer's disease and early-onset Alzheimer's disease (EOAD). We studied the role of the APOE*2 allele in the pathogenesis of EOAD in a Dutch population-based study of 175 probable EOAD patients with onset age at or before 65 years and 532 age-matched controls. In our population, there was no evidence for a protective effect of the APOE*2 allele on the risk of EOAD. When pooling the APOE2E2, APOE2E3 and APOE2E4 genotypes, the relative risk associated with the APOE*2 allele was 1.2 (95% confidence interval: 0.7-2.0). However, our data show that among EOAD patients survival for APOE*2 carriers was significantly reduced. When restricting the analysis to patients ascertained early after diagnosis at a stage of disease when mortality is low, our data suggest an increased risk of EOAD for subjects with APOE2E2, APOE2E3, APOE3E4 and APOE4E4 genotypes.

7

Genetic epidemiologic differences in p16+ and p16- melanoma-prone families.
AM Goldstein, MC Fraser, JP Struewing, CJ Hussussian, NC Dracopoli, WH Clark, Jr.*, MA Tucker. Natl Inst of Health, Bethesda, MD; *Harvard Medical School, Boston, MA.

p16/CDKN2, a chromosome 9p candidate gene for melanoma (CMM), was previously examined in 19 CMM kindreds. Ten kindreds had mutations that co-segregated with disease (p16+) whereas the other 9 kindreds did not (p16-). We looked for clinical and genetic epidemiologic differences between these two groups.

There were no differences in age at CMM diagnosis (median 31 vs 35.5, p=0.32), number of CMM tumors (median 1.0 vs. 1.0, p=0.24), or tumor thickness (median 0.79 vs. 0.80, p=0.78), between p16+ and p16- kindreds. The risk of CMM was similarly increased more than 60-fold in both groups. The most striking difference occurred for risk of other tumors. The prospective risk of pancreatic cancer was 13-fold increased (O/E = 2/.154 = 13.0, 95% CI 1.5-46.8) in p16+ kindreds. In contrast, there were no cases of pancreatic cancer in p16- kindreds. Analysis of the prospective and retrospective periods revealed that 5 of 10 p16+ kindreds had >=1 case of pancreatic cancer versus 0 of 9 p16- kindreds. No other tumors showed differences between the groups.

Previous examinations of familial melanoma have inconsistently shown relationships between melanoma and pancreatic cancer. p16 status may help differentiate CMM types because pancreatic cancer appears associated with only p16+ familial melanoma. The findings presented here suggest that incorporation of genetic factors may help resolve the heterogeneity (and controversy) surrounding familial melanoma.

8

SEGREGATION ANALYSIS OF CUTANEOUS MALIGNANT MELANOMA IN QUEENSLAND, AUSTRALIA

Joanne F. Aitken,* Joan E. Bailey-Wilson, Robert C. Elston, Adele C. Green, Robert MacLennan, Nicholas G. Martin (Queensland Institute of Medical Research, P.O. Royal Brisbane Hospital, Queensland, Australia 4029)

The genetic epidemiology of cutaneous melanoma was investigated in 9,726 individuals in 1,911 pedigrees ascertained through a population-based sample of histologically confirmed cutaneous melanoma cases diagnosed in Queensland, Australia, between 1982 and 1990. Information on melanoma history and melanoma risk factors was obtained by mailed self-administered questionnaire from all relatives. Relatives' reported melanomas were checked in medical records, and only those relatives for whom histological confirmation was obtained (n = 430) were counted as having the disease in theanalysis. Segregation analysis of melanoma was performed using the REGTL program in the S.A.G.E.computer package to examine compatibility of melanoma occurrence in these families with Mendelian segregation of an autosomal gene. We found no evidence to support the hypothesis of major gene inheritance. The inclusion of regressive familial effects and relatives' melanoma risk factors (ability to suntan, skin colour, naevus

density) significantly improved the fit of all models. The lack of fit of the genetic model may be due to the presence of other important risk factors shared by family members, in particular, sun exposure.

9

Sib-Pair Linkage Analysis using two loosely-linked marker loci

P.Holmans

(Dept. of Psychological Medicine,
Univ. of Wales College of Medicine, Cardiff)

It is well known that the power of all linkage analysis methods, including affected sib-pair analysis, depends greatly on the polymorphism of the marker. If there is no highly-polymorphic locus in the area under study, it may be possible to combine two or more loci together to form haplotypes, provided they are tightly linked. This introduces possible phase uncertainties, but the amount of power lost due to this is small (Holmans & Clayton 1995, submitted).

A more serious problem with haplotype-building occurs when the marker loci are not tightly linked, since families which exhibit recombination cannot be used. This paper attempts to deal with this problem by introducing a two marker locus extension to the likelihood-ratio method of affected sib-pair analysis originally proposed by Risch (1990, Am J Hum Genet 46:242-253)

The method involves dividing the interval between the marker loci into a number of equally-sized segments, to form a grid. At each gridpoint, the likelihood of the data can be expressed as a linear combination of the ibd sharing probabilities of the affected pair at a hypothetical disease locus situated at that point. The likelihood can then be maximised with respect to the ibd probabilities, subject to the restrictions recommended by Holmans (1993, Am. J. Hum. Genet 52:362-374) to give a maximised likelihood ratio. This is done for each gridpoint in turn, the largest of the likelihood ratios being taken as the test statistic for linkage.

Approximate test criteria were found by simulation, and the power of the method was compared to that obtained when only one of the marker loci was used. The increase in power gained by the two-locus method is greatest when the loci are not very polymorphic. Using the two-locus method can reduce the necessary sample size by as much as 40%. The increase in power gained by typing parents was investigated. The conclusions reached were similar to those obtained for one-locus analysis by Holmans (1993). Typing both parents reduces the sample size by at most one-third and therefore requires more individuals to be typed to achieve the same power.

10

Survival Analysis "Residuals" for Age-Adjustment in Sib-Pair Linkage Analysis.

RL Hanson, WC Knowler. NIDDK, Phoenix, AZ, USA

Sib-pair methods of linkage analysis may be useful in detecting genetic factors involved in chronic diseases. As many such disorders are strongly age-dependent, we propose a method of accounting for this in linkage analysis. Marker data were simulated to investigate the properties of this approach, using a survival analysis "residual", for detecting diabetes susceptibility genes in 1252 sib-pairs from 226 Pima Indian nuclear families.

For various values of marker heterozygosity (h) and recombination fraction (θ) between disease and marker, simulations were conducted under a model, from previous segregation analysis, where the disease gene influences age at onset. Diabetes was defined as a dichotomous variable "Y" (0 if unaffected, 1 if affected) and "X" was the age at examination if unnaffected or age of onset if affected. The survival analysis residual (R) was defined as: $R = Y - CI_X$, where CI_X represents the cumulative incidence at age X. In this analysis CI_X was calculated by the Kaplan-Meier method from a longitudinal study of incidence of diabetes in the same population. The Haseman-Elston test for linkage to a quantitative trait was then applied to R.

For replicates at $\theta = 0.5$, there was no appreciable increase in rate of Type I error using R as the trait in linkage analysis. Use of the survival analysis residual resulted in a modest increase in power over comparison of concordant and discordant pairs and a more substantial increase over analysis of affected pairs. For 500 replicates at $h = 0.8$ and $\theta = 0.025$, power to detect linkage ($p < 0.05$) was 52 % for the affected-only analysis, 68 % for comparison of concordant with discordant pairs, and 71 % for the survival analysis method. An alternate approach to age adjustment, use of the population age of onset distribution to estimate the probability that the sib-pair is discordant in susceptibility, reduced the power to detect linkage in this sample.

A survival analysis method can enhance the power of sib-pair linkage analysis. This method would presumably be most useful when analyzing families with variation in age or when the survival analysis incorporates environmental covariates.

11

Identifiability of Segregation and Ascertainment Parameters Using GEE When Families Are Ascertained Through A Child Having An Extreme Phenotype. J.S. Grove, University of Hawaii, Honolulu, Hawaii. U.S.A.

Segregation analysis using the method of maximum likelihood requires the assumption of a specific joint likelihood for the pedigree data. Generalized Estimating Equations (GEE) is an alternate method of estimation which does not require such strong assumptions. The segregation parameters for the general model (two alleles, three genotypes, plus polygenic background) have been shown to be identifiable from nuclear families using GEE, although third order crossproduct moments are required.

If nuclear families are ascertained through a single proband having a phenotype value greater than a specified cutpoint, three new parameters are introduced, the ascertainment probabilities for the three genotypes. This almost doubles the number of parameters to be estimated and poses a potentially serious problem. It is shown that including a random sample of unrelated individuals to the ascertained families provides sufficient information for identifiability. In addition, if ascertainment is for phenotypes extreme enough that only one homozygote is ascertained, the basic segregation parameters are identifiable using only low order moments.

12

MODELLING FAMILIAL CORRELATIONS USING MULTILEVEL MODELLING. PR Burton', L. Gurrin. (Division of Biostatistics and Computing, Institute for Child Health Research, Perth, Australia)

Advances in biotechnology have placed an emphasis on the collection of data from families rather than individuals. However the implications that intra-family correlations have for analysis are not always understood by the scientists who work with these data. Thus, data may be collected from a series of nuclear families in order to quantify a fixed effect of interest; *e.g.*, the

association between a putative mutation and a quantitative phenotype. In this setting, the correlations between siblings, between parents and children and between parents are a nuisance which must be addressed appropriately. Although this may be done using generalized estimating equations (*e.g. GEE4*), or mixed modelling (*e.g. Fisher*), most such approaches demand an understanding of a programming language and are difficult to use without experience. This difficulty is appropriate if interest centres upon the correlation structure itself. However, if the correlation is merely a nuisance which must be addressed in some way more reasonable than simply assuming *independence* the difficulty may be unwarranted. We have devised a way to analyse such data using multilevel modelling in *ML3*. *ML3* supports hierarchical mixed modelling in an environment (based on *Minitab* via *Nanostat*) in which it is easy to work. *ML3* fits nested covariance structures with up to 3 levels, when a response is Multivariate Normal. Nuclear families generate non-nested correlations but these may be parameterised as a 3-level nested structure. An *ML3* macro was written to transform estimated random effects to correlations. Standard errors were estimated using the delta method. Simulations showed the model to provide consistent estimates of fixed effects, correlations and standard errors which were similar to those generated by *Fisher* and *GEE4*. *ML3* provides non-statisticians with a friendly environment in which to store, manipulate and analyse data. By fitting a standard covariance structure and using the *ML3* macro it is easy to fit models which permit the estimation of one or more fixed effects of interest while adjusting appropriately for the correlation structure. We will discuss model construction and interpretation, and present results from real and simulated examples.

13

Interval mapping of human quantitative trait loci using data from nuclear and extended families.

J.M. Olson, Department of Biomathematics, Roswell Park Cancer Institute, Buffalo, New York.

The use of multiple markers, rather than a single marker, can increase the likelihood of detecting linkage to a locus underlying a quantitative trait. Single-marker sib-pair linkage methods have recently been extended to the multiple marker case (e.g., Olson, 1995) for independent sib pairs. These methods first estimate, using likelihood methods, the joint distribution of marker identity-by-descent (IBD) sharing and then apply the Haseman-Elston regression method in an interval mapping framework. This paper discusses the use of these methods when pairs are not independent, with emphasis on the estimation of the joint distribution of marker IBD sharing. An extension of the algorithms of Amos et al. (1989) provide a means to compute multipoint IBD states from sibships with and without missing parental marker information. The correlations among sib pairs are treated using generalized estimating equations approach as in the single-marker case (Olson and Wijsman, 1993). When extended to relative pairs in a set of extended pedigrees, these methods potentially provide a most powerful means for the robust mapping of quantitative trait loci by including information from all available relative pairs and multiple marker loci. Simulations are used to examine the size, power and robustness of the approach.

14

Approximating Multipoint Lod Scores.

V. Haghighi, J. Ott. Columbia University, NY.

Linkage analysis between disease and multiple marker loci is computationally intensive, especially for polymorphic markers (i.e. heterozygosity ≥ 70%). We propose an approximation method which is expected to be particularly useful in presence of untyped individuals. Specifically, we apply a type of importance sampling to multipoint linkage analysis.

We will provide a brief summary of this method for the case of three loci , yet it can be easily extended to any number of loci. In this case one locus is the disease and the other two loci are markers. Let x, y and z denote the phenotype arrays at three loci for all members of a given family. The likelihood of the data under the parameter vector θ, (i.e. the recombination fractions) is proportional to the probability of the observations summed over all possible genotype arrays f, g and h for all family members at each of the three loci. The likelihood may be written as:

$$p\left(x,y,z;\theta\right) = \sum_{f}\sum_{g}\sum_{h} p(x,y,z|f,g,h)\,p\left(f,g,h;\theta\right)$$

Given that the penetrances depend only on individual loci, we can rewrite the above likelihood as:

$$p\left(x,y,z;\theta\right) = p(x)p(y)p(z)\sum_{f}\sum_{g}\sum_{h}\left(\frac{p\left(f,g,h;\theta\right)}{p(f)p(g)p(h)}\right)p(f|x)p(g|y)p(h|z)$$

The P(x), P(y) and P(z) are single-locus phenotypic likelihoods. The P(f), P(g) and P(h) are the mendelian genotypic probabilities for each locus. The P(f | x), P(g | y) and P(h | z) define the single-locus sampling distributions, from which we simulate replicates via random sampling. We prefer working with the single-locus sampling distributions over the multi-locus one, due to the gain in computational efficiency in simulating single-locus genotype arrays. To compensate for any biases resulting from this sampling scheme , we replace the joint likelihood, P(f, g, h ; θ) by the ratio, P(f ,g, h ; θ) / P(f)P(g)P(h). Note that the only multipoint quantity to be evaluated is P(f, g, h ; θ), and this is calculated very quickly, since the genotypes of untyped individuals have been simulated using the above sampling distributions. The end result is that this approximation to the multipoint lod score can be arrived at in a fraction of the time required for a full likelihood calculation. Indeed, this method can be used as a precursor to fine mapping.

15

Segregation Analysis of Breast Cancer in France: Search for Interactions between Genetic and Reproductive Risk Factors.

N. Andrieu[1] and F. Demenais[2].
[1]INSERM U351, Villejuif, France, [2]INSERM U358, Paris, France.

In order to better understand the role of genetic and reproductive factors in breast cancer (BC), family data were systematically collected in two French hospitals (Instituts Gustave Roussy and Curie) between 1987 and 1989. Information on BC status and reproductive factors was recorded among 288 probands and their female first-degree relatives. Segregation analyses were conducted using the class D logistic regressive model wich accounts for a variable age at onset of the disease (Abel & Bonney, 1990, Genet Epidemiol, 7:391-407), as implemented in the program REGRESS (Demenais & Lathrop, 1994, Genet Epidemiol, 11: 291). When analyses were peformed ignoring the reproductive factors, our results indicated segregation of a rare autosomal dominant gene (q=0.0006) with penetrance increasing with age differently in susceptible and non-susceptible individuals (gene x age interaction). The lifetime risk for gene carriers is 0.69 and the proportion of sporadic cases is 0.91. When the reproductive factors are included as covariates in the model, the evidence for a mendelian transmission of the major gene is stronger. A higher penetrance is reached at a younger age in gene carriers who are also at risk for reproductive factors. Tests of interaction

between the major gene and each of the following risk factors did not reach statistical significance: age at menarche, number of children, experience of abortion and menopausal status. However, the effects of age at menarche and number of children on BC risk appear different in susceptible as opposed to non-susceptible individuals.
This study outlines the importance of accounting for epidemiological risk factors in familial analyses in order to better assess the effect of genetic factors and better predict the risk of disease.

16

Polymorphisms in carcinogen-metabolizing genes and breast cancer susceptibility. C.B. Ambrosone[1], J.L. Freudenheim[1], S. Graham[1], J.E. Vena[1], J.R. Brasure[1], A.M. Harrington[2], T. Ford[2], and P.G. Shields.[2] [1]Dept. of Social & Prev. Med., SUNY at Buffalo, Buffalo, NY. [2]National Cancer Institute, Bethesda, MD.

N-acetyltransferase (*NAT2*) and glutathione S-transferase (GSTM1) are involved in metabolism of carcinogens, including those in cigarette smoke. A genetic polymorphism in either may result in decreased detoxification of carcinogens. In a case-control study of 159 postmenopausal women with incident, primary, histologically confirmed breast cancer and 203 controls, matched on age and county of residence, we assessed the interaction of polymorphisms in *NAT2* and GSTM1 with cigarette smoking and breast cancer risk. PCR-RFLP analyses were performed. GSTM1 did not modify the association between smoking and risk. Although neither smoking nor *NAT2* genotype were independently associated with breast cancer risk, women with the slow acetylation *NAT2* genotype (56%) were at increased, dose-dependent risk from smoking ($p<0.001$). Smoking did not affect risk for rapid acetylators. Among slow acetylators, smoking ten years prior to the interview increased risk of breast cancer in comparison to non-smokers (4th quartile adjusted odds ratio, 7.20, 95% confidence interval, 2.25-23.02). These data suggest that, for women with the slow *NAT2* genotype, smoking may be a strong risk factor for breast cancer. Furthermore, evaluation of genetic variability in metabolism of carcinogens, as well as exposure assessment in epidemiologic studies, may explain previously unclear associations between environmental factors and breast cancer risk.

17

Genetic epidemiology of *Loa loa* filariasis in Cameroon.
A. Garcia[12], L. Abel[2], M. Cot[1], S. Ranque[3], P. Richard[3], M. Boussinesq[3], JP. Chippaux[3].
[1]ORSTOM OCEAC, Yaounde, Cameroon, [2]INSERM U436, Paris, France and [3]ORSTOM Centre Pasteur, Yaounde, Cameroon.

Loa loa filariasis, a parasitic disease transmitted by the bite of an insect vector, is found only in the tropical rain forest of central Africa where more than 30 million people live. Severe complications as endocardial fibrosis or renal diseases have been reported. Within hyperendemic regions exposure to *Loa loa* may approach 100%, with approximatively 30% of the population being microfilaremic (i.e. carrying microfilariae (MF) in their blood). To explain this relatively low rate of MF carriers the existence of a genetic control of the outcome of filarial infection have been put forward.
A familial longitudinal one-year survey of *Loa loa* infection was carried out in an endemic area of Southern Cameroon between april 1992 and april 1993. The follow-up concerned the total population of a forest village (n=667) with complete selection of all the 45 families that can be constructed from these individuals. Parasitologic samplings were performed every two months to assess individual parasite densities (PD) and their variation over time. The first step was to study the variation of the individual microfilarial status (IMS) over time at both the qualitative (i.e. microfilaremic/nonmicrofilaremic) and the quantitative (PD) level, and to determine the measured covariates that influence this IMS. The results show clearly the high stability of the IMS (qualitative and quantitative) over time. MF carriers were generally the same individuals over time and the variability of their PD was very low. Among 128 subjects sampled 6 times, 61% were never MF carriers whereas 13% were expected under the hypothesis of random occurence of negative PD within individuals ($p<10^{-9}$). Furthermore, age was the relevant factor that influenced the individual microfilarial status in the whole population, but not the level of microfilaremia in the population of microfilaremic individuals. These results are consistent with the hypothesis of an individual susceptibility to the outcome of filarial infection and suggest that genetic factors might be involved in host defense mechanisms against loiasis infection. The second step is a segregation analysis of the IMS using regressive models and preliminary results seem to confirm the genetic hypothesis.

18

Vitamin intake as a modifier of the association between smoking-related lung cancer and GSTM1 deletion. M. Garcia-Closas, K.T. Kelsey, J.K. Wiencke, D.C. Christiani (Harvard School of Public Health, Massachusetts General Hospital, Boston, MA 02115, and University of California at San Francisco CA 94143)

Low intake of antioxidants and GSTM1 deletion have independently been associated with smoking related lung cancer. Because both factors may modify oxidative damage within the lung, we studied their potential interaction in an on-going case-control study with 169 cases and 174 controls. Fifty percent of the cases were adenocarcinoma, 35% squamous cell carcinoma (SCC), and 16% other tumors. Nutrient intake was measured with a food frequency questionnaire and was used to classify individuals into group of high and low antioxidant vitamin intake. For all tumors combined, there was evidence for effect modification by vitamin C intake (OR=2.1, 95%CI 0.9-4.7 for low intake, OR=0.8, 95%CI 0.5-1.3 for higher intake, p-value for test for interaction 0.04), and vitamin A intake (OR=2.5, 95%CI 1.1-5.7 for low intake, OR=0.7, 95%CI 0.4-1.3 for higher intake p-value for test for interaction 0.02). When considering histologic types separately, this effect appeared to be present only for SCC (p-values for test for interaction 0.01 for vitamin C and 0.09 for vitamin A) but not for adenocarcinoma, although in this preliminary analysis the numbers were too small to give us accurate estimates of effect by tumor type. Overall, these results suggest that GSTM1 deletion increases the risk of SCC among subjects with low intake of antioxidant vitamins.
(Supported by P01 ES06409)

19

Evidence for a major gene influencing observed increases with age in diastolic blood pressure levels.

L S-C Cheng[1], D Carmelli[1], SC Hunt[2], RR Williams[2]
[1] Health Sciences Program, SRI International, Menlo Park, CA; [2]Dep. of Internal Medicine, U. of Utah, Salt Lake City, UT, U.S.A.

The contribution of genetic factors to blood pressure (BP) levels is well established. The multifactorial nature and heterogeneity of this trait often preclude the search for major gene effects. The observation that BP increases with age implies at least, or both, a cardiovascular natural response to biologically accumulated stimuli, and/or a significant role for age-gene-environment influences. The purpose of the present study was to investigate the mode of inheritance of diastolic BP (DBP) as a function of age. We hypothesized that a major gene, in addition to many other genetic and environmental factors, controls the observed increase with age in diastolic BP levels. To avoid confounding due to maturation, analyses were restricted to adults only (age >=18 y/o). After exclusion of subjects on antihypertension medications, the total sample size available was 965 individuals in 73 pedigrees collected in Utah as part of a longitudinal cardiovascular family study. Over an average of 7.2 years of follow-up we observed a 3.18% statistically significant change from baseline in diastolic BP. Test-re-test correlations for DBP within individuals were 0.56 in males and 0.58 in females. Familial correlations for DBP change scores, weighted on pedigree size and adjusted for age were 0.01, 0.04, and 0.16 for spouses, parent-offspring, and siblings, respectively. A segregation analysis of DBP change scores suggest significant age effects, but no sex effects. The most parsimonious model identified a major recessive gene effect with a gene frequency of p = 0.22. The genotypic mean of DBP change for the susceptible subgroup was 30.47% with a 4.1% modifier effect due to age, compared to 1.79% and a -1.2% age effect for the non-susceptible genotypes. Two additional pedigree analyses were conducted on DBP levels at baseline and follow-up. In these analyses we found significant sex and age effects and a moderate heritability (20-30%) on DBP levels at both visits. The best fitting model for baseline levels was a mixed codominant Mendelian model with a gene frequency p = 0.22. Analyses of the follow-up data, found that genetic effect remained significant but transmission parameters deviated from the Mendelian segregation ratios. The present study demonstrates the utility of longitudinal family data to detect major gene effects possibly involved in changes with age of BP levels.

20

A bivariate segregation analysis of fat mass and body mass index reveals a common major gene for both traits.

A.G. Comuzzie[1], B.D. Mitchell[1], J. Blangero[1], L.D. Atwood[1], M.C. Mahaney[1], M.P. Stern[2], J.W. MacCluer[1].
[1]Southwest Foundation for Biomedical Research, and [2]University of Texas Health Science Center, San Antonio, Texas, U.S.A.

We previously reported a major gene for fat mass (FM) in Mexican Americans and have replicated this finding in another sample here. An earlier study also suggested evidence of a major gene influencing the body mass index (BMI) in this population. Here we employ bivariate segregation analysis, using a modified version of PAP, to determine if the major gene for FM also influences BMI. Data from 465 individuals in 31 Mexican American families participating in the San Antonio Family Diabetes Study were used in this work.

A polygenic analysis found significant genetic and environmental correlations between FM and BMI (0.87 and 0.94, respectively). The hypothesis that the high genetic correlation between these two traits is due to the effect of a single major gene was tested by comparing two models: (1) a model in which FM and BMI were influenced by a single major gene and also by shared polygenes, and (2) a nested model with a major gene that influenced only FM, in which all of the genetic correlation was through shared polygenes. The second model was strongly rejected ($p < 0.00001$), indicating that the major gene for fat mass exerts a pleiotropic effect on BMI. These findings show that bivariate analysis of correlated traits can strengthen evidence for a major gene. Supported by NIH grants GM15803, DK42273, and HL45522.

21

An Association between Variants of the ACE Gene and Obesity in Men

Richard Cooper, Terrence Forrester, Norma Anderson, Ryk Ward; Loyola Medical School, Maywood, IL, University of the West Indies, Mona, Jamaica, University of Utah, Salt Lake City, UT.

ACE is a ubiquitous enzyme with multiple physiologic roles. An impact of genetic variants at the ACE gene has been noted on both blood pressure (BP) and left ventricular hypertrophy. In a population survey in Jamaica we detected a relationship between the ACE insertion allele of the insertion/deletion (I/D) polymorphism, ACE activity, and body mass index (BMI) and skinfold thickness. First, ACE activity levels varied by genotype: II = 135, ID = 158. DD = 182 ($p < .001$). These data are consonant with previous findings among whites, and suggest segregation at the ACE locus accounts for a significant fraction of plasma ACE variability in populations of African descent. Second, obesity was related to genotype among the men (n = 171): BMI, II = 25.5, ID = 24.2, DD = 23.1 ($p < .02$); subscapular skinfold, II = 17.1, ID = 17.1, DD = 14.5 ($p < .03$). A similar trend was noted in BP; sys BP, II = 131, DD = 125. Sexual dimorphism in obesity could result in heterogenous effects across gender, where the women were more obese (mean BMI = 27, n = 326). This effect might also be mediated by testicular ACE through steroid hormone regulation. These findings, to our knowledge, represent the first demonstration of a candidate gene for obesity, and may help explain the interconnection between BMI and BP and gender differences in hypertension risk.

22

Identification at the amino acid level of the HLA component to IDDM. G. Thomson, A.M. Valdes, H. Salamon, T. Jorma, K. Ronningen. Universities of California, Berkeley, CA 94720-3140, Helsinki and Oslo.

Current molecular models of insulin dependent diabetes mellitus (IDDM), Asp at DQB1 57 protective and Arg at DQA1 52 predisposing, do not explain the known genetic heterogeneity of IDDM. We have developed two independent methods to identify the amino acids in the HLA region involved in IDDM. The "unique combinations" computer algorithm identifies all single and multisite amino acid

differences between specified sets of alleles, haplotypes and genotypes from patients and controls. The "haplotype method" is a statistical test of whether all relevant amino acids involved in disease have been identified. For haplotype (chromosome) combinations containing all the amino acid sites involved in the disease process, the relative frequencies of amino acid variants at sites not involved in disease are expected to be the same in patients and controls. Our results prove that the combination DQA1 52 - DQB1 57 does not include all the predisposing elements. Specific amino acids have been identified which account for the predisposing effect in both Japanese and Caucasians. Other ethnic groups are also being studied, as well as HLA-DRB1 and DPB1 variation.

23

Family study of lipoprotein lipase gene polymorphisms and plasma triglyceride levels

Jean-Louis Georges[1], Anne Régis-Bailly[2], Driss Salah[2], Roger Rakotovao[1], Gérard Siest[2], Sophie Visvikis[2], Laurence Tiret[1].
[1]INSERM U258, Paris, and [2]Centre de Médecine Préventive, Vandoeuvre-lès-Nancy, France.

To progress in the understanding of the role of the lipoprotein lipase (LPL) gene in the determination of the plasma triglyceride (TG) concentration , a study was performed in 193 healthy nuclear families (785 subjects, including 399 offspring), volunteering to have a free heath-checkup examination. The pattern of familial correlations was compatible with no spouse and father-offspring resemblance, and significant mother-offspring ($r \simeq .24 \pm .05$) and sib-sib ($r \simeq .30 \pm .06$) correlations. Association between TG concentrations and two RFLPs of the LPL gene (*HindIII* and *PvuII*) was first investigated by a familial measured genotype analysis, specifying sex- and age-dependent gene effects. The effect associated with both polymorphisms was significant only in fathers, the H+ and P+ alleles being associated with raised levels of TG.

The hypothesis of a major gene in linkage disequilibrium with the RFLPs was then tested by combined segregation and linkage analysis. This analysis did not provide clear-cut results, but mothers and offspring were poorly informative in this analysis because of the lack of polymorphism effect among them. For this reason, a commingling analysis conditional on the marker genotype was performed in the subset of fathers. The results of this analysis supported the hypothesis of a major gene whose frequency was estimated as $.304 \pm .043$, and being in linkage disequilibrium with H+ ($p < .005$) and P+ ($p < .05$). This major gene would explain 39-58% of the variability of the trait in fathers, according to the model considered.

These results suggest that the LPL gene is involved in the determination of triglyceride levels through the action of a common functional polymorphism in linkage disequilibrium with *HindIII* and *PvuII*. The effect detectable only in fathers indicate a likely modulation of the LPL expression by hormonal or lifestyle factors.

24

Multivariate associations of ACE and AGT variants with coronary heart disease and hypertension. I.B. Borecki[1], M. Higgins[2], J.M. Lalouel[3], G. Heiss[4], and the Family Heart Study investigators. [1]Division of Biostatistics, Washington Univ. School Medicine, St. Louis; [2]NHLBI, Bethesda; [3]Howard Hughes Medical Institute, Univ. Utah, Salt Lake City; [4]Dept. Epidemiology, Univ. North Carolina, Chapel Hill.

The insertion (I) / deletion (D) polymorphism in the angiotensin converting enzyme (ACE) locus and the M235T substitution (alleles 'M' and 'T') in the angiotensinogen (AGT) locus have been associated with myocardial infarction (MI) and hypertension, respectively. In an attempt to replicate these findings, subjects participating in the *F*amily *H*eart *S*tudy (FHS) were studied using a case-control design for tests of association; allele frequencies for both loci were determined in random samples. CHD cases (N=318) were defined as subjects with validated MI, silent MI, PTCA, or CABG; an equal number of controls were matched within 5 years of age, by sex, and were free of CHD themselves, and in any parent or sibling, and did not use lipid lowering medications. Severe hypertensive cases (N=376) used 2 or more anti-hypertensive medications and reported at least one parent with high blood pressure (BP); controls were matched on age and sex, and were required to be normotensive, using no antihypertensive medications, and had no high BP in parents. Multivariate logistic analysis was used in each case including, in addition to ACE and AGT, measurements of BP, lipids, lipoproteins, and apolipoproteins, BMI, and cigarette usage as potential independent predictors. For the CHD analysis, hypertensive status, cholesterol, LDL, HDL, apo AI, Lp(a) and cigarette usage were significant in addition to AGT 'TT' (OR=1.84, 1.09-3.10 95% CI). By restricting the sample to low risk subjects (those below median for both apo B and BMI), only LDL and triglycerides were significant risk factors, along with a strong interaction between the DD and TT genotypes (OR=5.02, 1.22-20.55), suggesting a complex interaction among risk factors and genotypes. For the hypertension analysis, significant predictors included CHD case status, BMI, HDL, triglycerides, as well as AGT genotypes; however, the effect of the latter was sex-dependent (P=.011). The effect of the TT genotype was more pronounced in females (OR=1.90) than in males (OR=1.52), with some dosage effect in heterozygotes. It appears that both these loci are significant predictors of CHD, and AGT of hypertension, independent of other established risk factors.

25

HLA-DQ Haplotypes and Susceptibility to Autoimmune Thyroid Disease (ATD) in Insulin-dependent Diabetes Mellitus (IDDM) Families. E. McCanlies*, J.S. Dorman, B. McCarthy, M.K. Kramer, J. Swan, J. Burke, L. O'Leary, A. Koehler, M. Trucco, T. Foley. University of Pittsburgh, Pittsburgh, PA.

Individuals with IDDM and their relatives frequently develop ATD, including Hashimoto's thyroiditis (HT) and Graves disease (GD). The familial clustering of these disorders may be due, in part, to shared HLA susceptibility genes that contribute to multiple autoimmune disorders. The Familial Autoimmune and Diabetes (FAD) Study is evaluating potential genetic / environmental determinants of the clustering of ATD among probands from the Children's Hospital of Pittsburgh IDDM Registry for 1950-65 and their family members. The presence of ATD is being assessed by high titres (>10 U/ml) of thyroid peroxidase and thyroglobulin antibodies (for HT) or TSH receptor antibodies (for GD), the level of TSH and T3, and a positive medical history or physical exam. To date, 162 IDDM probands, 157 siblings and 163 parents have participated. Twenty-two parent-offspring HT families have been identified. Four parental HLA-DQA1 and DQB1 haplotypes were defined in 9 of these families. Comparisons of the 27 haplotypes in HT individuals with the 9 non-HT haplotypes revealed a 2-fold increase in DQA1*0501-DQB1*0201 (22% vs. 11%) and a decrease in DQA1*0301-DQB1*0302 (22% vs. 33%). Five of the 9 HT parents carried DQA1*0501-DQB1*0201, 4 of whom transmitted this haplotype to their HT offspring. Only 2 HT parents carried DQA1*0301-DQB1*0302; 1

transmitted it to an HT offspring. Graves disease was rare in this cohort, found in 5 individuals (4 were mothers). Two of the 3 mothers who were typed carried DQA1*0301-DQB1*0302. DQA1*0501-DQB1*0201 was not detected in a Graves patient. These results suggest that susceptibility to HT may be related to DQA1*0501-DQB1*0201, while GD may be associated with DQA1*0301-DQB1*0302. These haplotypes are strongly diabetogenic in Caucasians and may contribute to the clustering of ATD in IDDM families.

26

Racial Differences in the Molecular Epidemiology of Insulin-dependent Diabetes Mellitus (IDDM) in the US. B.J. McCarthy*, J. Roseman, E. McCanlies, C.S. Ivie, A. Rahman, M. Trucco, J. Dorman, and the WHO DiaMond Molecular Epidemiology Sub-Project Group. University of Pittsburgh, PA and University of Alabama at Birmingham, AL.

In Allegheny County, PA, and Jefferson County, AL, the incidence of IDDM among Whites is higher than that for Blacks (PA: 17 vs. 11/100,000/yr; AL:15 vs. 7/100,000/yr,respectively). Differences in risk may be due to racial or population variation in IDDM susceptibility genes. As part of the WHO DiaMond Molecular IDDM Epidemiology Sub-Project, HLA-DQA1 and DQB1 typing is being performed for IDDM cases (n=44,59,Whites; n=33,29,Blacks) and non-diabetic controls (n=90,61, Whites; n=40,53, Blacks) from PA and AL, respectively. DQB1 alleles coding for an amino acid other than aspartate in position 57 (ND), particularly DQB1*0302, were significantly associated with IDDM in all groups. DQA1*0301 (coding for arginine (R) in position 52) was also consistently associated with IDDM. Homozygotes for DQB1*ND or DQA1*R had significantly increased risks relative to DQB1*D or DQA1*NR homozygotes, respectively, in all groups. Double homozygotes (DH) were at highest IDDM risk. Genotype distributions for PA and AL were compared. White IDDMs were similar (i.e.,41% vs. 40% from PA and AL, respectively, were DH). No significant differences were observed for PA vs. AL Black IDDMs. Controls were also similar. Among Whites, 9% vs. 3% from PA and AL, respectively, were DH. Proportions for Blacks were 0% and 8%, respectively. Although more PA White than PA Black controls were DH (9% vs. 0%, $p<.05$), no racial differences in genotype were observed for controls from AL. These data suggest that variation in the DQA1*R-DQB1*ND genotype distribution may contribute to racial differences in IDDM incidence within a population, but are unlikely to explain small risk differences across populations of the same race.

27

Familial correlations of anthropometric variables and relative fat distribution among African Americans
C.N. Rotimi, R.S. Cooper; Loyola University Stritch School of Medicine, Maywood, IL, U.S.A.

Obesity has emerged as a major public health problem in the African-American community and its determinants are poorly understood. The purpose of this study was to examine familial patterns of the degree of heaviness and of regional fat deposits. Seven anthropometric measurements (height, weight, body mass index [BMI], waist and hip circumferences, waist-to-hip ratio and arm circumferences) were examined in a population sample of African-American nuclear families recruited from the Chicago area. Participants included 162 parent pairs and 114 sets of siblings (a total of 295 sons and daughters). Adjustments to correct for the effect of age and sex were performed separately for parents and siblings and the residuals were transformed to normalized z-scores to calculate intraclass correlations between family members. With the exception of waist-to-hip ratio, all variables showed positive familial resemblance, although the degree and level of statistical significance varied by type of relationship. Significant familial resemblance was observed for peripheral fat deposit, as measured by hip and arm circumferences (range: 0.04 for parents to 0.42 for daughters), and obesity, as measured by BMI (range: 0.12 for parents to 0.33 for mother-daughter comparison). Significant sex and generation differences were observed for height and hip. We conclude that African Americans exhibit significant familial resemblance for stature, degree of heaviness and regional fat deposits. Understanding the relative contributions of shared genes and shared environmental factors to the observed familial patterns could yield important insights into the origins of obesity in this high risk population.

28

Genetic and environmental determinants of sustained intense and intermittent moderate physical activity levels in middle-aged male twins. D.S. Lauderdale[1], R. Fabsitz[2], V. Ramakrishnan[1,3], J. Goldberg[1,3].
[1]UIC-SPH Epid-Biostat; [2]NHLBI; [3]Vietnam Era Twin Registry, Hines, IL.

Physical activity level and fitness, assessed by diverse survey and clinical measurement methods, have been demonstrated to reduce the risk of mortality and chronic disease. New recommendations from the Centers for Disease Control and the American Academy of Sports Medicine suggest that moderate activities, may confer health benefits. This study estimates genetic and environmental determinants of both *sustained intense* and *intermittent moderate* forms of leisure-time physical activity. The study population is 3,344 pairs of male twins age 33 to 51 both of whom served in the military. These twins were asked eleven questions about regular participation in specific, intense, athletic activities and walking and stair climbing for exercise. For each of the eleven questions, MZ correlations were higher than DZ correlations, indicating genes play a role in the variation observed in the population. For regular performance of five sustained intense activities (running, bicycling, swimming, racquet and other sports) significant heritability estimates ranged from 0.38 to 0.59 and common environment was not significant. Heritability estimates for six intermittent moderate activities (all related to discretionary walking or stair climbing) ranged from 0.26 to 0.41. A factor analysis produced two factors accounting for 49% of the variation: an intense and a moderate activity factor. For the intense activity factor, heritability was 0.46, while it was 0.29 for the moderate factor. Common environmental effects were negligible. Adjustment by age and race did not affect the results.

29

Coronary artery disease risk factors may differ by apo E phenotype: The Oklahoma Marker Study. J.E. Eichner and W.E. Moore. University of Oklahoma, Oklahoma City, OK, U.S.A.

Epidemiologic studies of coronary artery disease (CAD) risk factors examine risk factors in the aggregate and generally show a modest impact on disease and require a large sample size. Classifying individuals by apolipoprotein (apo) E phenotype and examining risk factors within phenotypes may help explain some inconsistencies among risk factors. The following study is based on consecutive consenting coronary angiography patients from two Oklahoma hospitals. There were a total of 484 white male patients who donated blood and allowed use of their medical and catherization reports and had available data. Apo E gene frequencies for the entire group did not differ substantially from apo E gene frequencies from the Multiple Risk Factor Intervention Trial (MRFIT) ($\varepsilon3$=.78, $\varepsilon4$=.15, $\varepsilon2$=.07 vs. $\varepsilon3$=.79, $\varepsilon4$=.15, $\varepsilon2$=.06) [$\chi2$=.85, 2df, p=.33]; both groups being men at risk for coronary heart disease. For men with the most common phenotype apo E3/3 (n=297), age, triglycerides, and lipoprotein (a) [Lp(a)] were the only variables significantly associated with having at least one coronary artery lesion $\geq$ 25% when age, body mass index, diabetes, fibrinogen, low and high density lipoprotein cholesterol, triglycerides and Lp(a) were used in the logistic model [p$\leq$.01]. These data agree with a previously published report that triglycerides is an important risk factor for apo E 3/3 homozygotes who are the most frequently found apo E phenotype in the general population, as well as, among cardiology patients.

30

Survival in parents of patients with activated protein C resistance (factor V Leiden)

E.T.M. Hille, T. Koster, F.R. Rosendaal, J.P. Vandenbroucke, Dept. of Clinical Epidemiology and the Haemostasis and Thrombosis Research Centre, Leiden University Hospital, The Netherlands

A poor anticoagulant response to activated protein C (APC resistance) is a common autosomal inherited abnormality due to a mutation at the APC-cleavage site in the factor V gene (factor V Leiden). APC resistance is associated with a tendency to venous thrombosis in heterozygous individuals. Survival in APC resistant individuals can only be studied prospectively, which would take many years. The Family Tree Mortality Ratio (FTMR) method offers another approach to study survival of carriers retrospectively. At least one parent of each index case carries the mutation which he/she has passed on to his/her offspring. A deleterious effect of the mutation on survival would therefore be apparent in parents of probands. We compared mortality in parents of APC resistant patients (observed) with that of the Dutch population (expected) adjusted for sex, age and calendar period. The ratio of observed to expected is the standardized mortality ratio (SMR).
In the Leiden Thrombophilia Study 92 patients had the factor V Leiden mutation (8 homozygous). Using national and municipal registries, we retrieved the dates of birth and death for the 184 parents of all patients. Follow-up extended from the date of birth of their affected offspring to the end-of-study date. No excess mortality was found (SMR 85/86.7, 95% CI 0.8-1.2). The SMR did not significantly differ from unity in both men and women, all age groups or calender periods. Obviously, only half of the parents were themselves carriers of the defect. The inclusion of a substantial number of normal parents would have diluted any observed effect on mortality, but cannot explain the absence of an effect. We conclude that APC resistance does not greatly affect overall survival in heterozygous individuals.

31

No evidence for anticipation in 9q34-linked idiopathic torsion dystonia. L. Almasy[1], S. Bressman[2], D. deLeon[2], N. Risch[3]. [1]Dept. Genetics, Yale Univ., New Haven, CT, [2]Columbia Presbyterian Medical Center, New York, NY, [3]Dept. Genetics, Stanford Univ., Stanford, CA.

There has been some suggestion of anticipation, or increasing severity of disease with succeeding generations of affected individuals, in idiopathic torsion dystonia. LaBuda et. al.(1) found significant differences in age of onset (AO) between affected parent-child pairs, but could not exclude bias due to more severely affected individuals having reduced reproductive fitness.

In anticipation due to expanding triplet repeats, we expect that affected children of affected parents would have longer repeats and earlier mean AO than affected children of normal parents. No such difference is seen in 114 affected offspring of 49 Ashekenazi Jewish and 7 Non-Jewish families linked to the 9q34 DYT1 locus or carrying the known DYT1 founder haplotype(Table 1).

This analysis assumes that non-penetrance in unaffected obligate carrier parents is due to repeat size and not to other modifying genetic or environmental factors. Alternatively, we would also expect that children of parents with younger AO should have an earlier mean AO than children of parents with older AO. No such difference is seen when we divide the sample at the median parental AO(Table 1). We also examine the proportion of cases in which symptoms became generalized, another measure of severity, but find no significant differences.

Table 1: Measures of Severity for Affected Children.

	Parents			aff w/	aff w/	
	unaff	aff		AO<14	AO14+	
AO	12.7	10.7	p=.23[a]	8.7	12.8	p=.09[a]
% gen	67.5	64.5	p=.77[b]	68.8	60.0	p=.61[b]
n	83	31		16	15	

a=t-test for difference between means b=chi-sq w/1 df

(1) LaBuda et.al.(1993) "Genomic imprinting and anticipation in idiopathic torsion dystonia" Neurology 43:2040-2043.

32

Case-control design options for family studies of measured genes and gene-environment interactions. JS Witte, WJ Gauderman, DC Thomas. University of Southern California, Los Angeles CA, USA.

Case-control designs provide an attractive approach for estimating the effects of measured disease genes and gene-environment interactions. In undertaking such studies, one must

decide who to include; for example, one might use probands and their unaffected siblings as cases and controls, respectively. To evaluate the appropriateness of different case-control approaches, we examine — through discussion and simulation — the following design options: case-sibling; case-cousin; case-cohort, where controls are selected from the probands' entire cohort of family members; case-"pseudo sibling," where controls are selected from the set of possible genotypes that could have been inherited from parents (e.g., the TDT approach); and case-population, where controls are randomly drawn from the general population. For our simulation studies, we randomly generated disease status for each subject in 17-member pedigrees (3 generations), assuming a single Mendelian diallelic major gene. For each trial, we generated 300 case-control sets (specific to each design), using pedigrees obtained from the population by single ascertainment of diseased probands in the third generation. We performed 2,000 trials per study. We give results from one of these studies below. For this parameter combination, the case-control designs were relatively comparable, except that the cohort approach underestimated the major gene effect and had subnominal coverage, and that the sibling approach had slightly lower power. We will present additional simulation results (i.e., using other parameter combinations) for assessing measured genes, as well as for using case-control designs to evaluate gene-environment interactions that involve measured genes.

Simulation results when evaluating a major gene.

Design	$\hat{\beta}$	se($\hat{\beta}$)	Coverg†	Power
sibling	0.71	0.28	94.8	74.1
cousin	0.69	0.21	94.8	94.0
cohort	0.58	0.19	90.6	84.3
TDT	0.66	0.20	94.1	92.9
population	0.70	0.20	94.5	94.3

* True values: β= 0.69, pop. prevalence=.05, allele frequency=0.1.
† 95% Coverage.

33

The Search for Additional Alzheimer Disease (Ad) Genes. MA Pericak-Vance1, M Ter-minassian2, PM Conneally3, GW Small4, AM Saunders1, L Yamaoka1, PA Locke2, M Pritchard1, RE Tanzi2, JF Gusella2, AD Roses1, JL Haines2. 1). Duke University Medical Center, Durham 2). Massachusetts General Hospital, Boston 3.) Indiana University Medical Center, Indianapolis 4). University of California, Los Angeles.

Alzheimer disease (AD) is a complex inherited disorder with proven genetic heterogeneity. To date, genes on chromosomes 21 (APP) and 14 (not yet identified) play a causative role in early-onset (<60years) familial AD, while the APOE gene on chromosome 19 is associated with both late onset familial and sporadic AD and possibly some forms of early onset AD. Although these genes account for the majority of AD, many familial cases are clearly not related to any of them. For example, in late onset AD, Risch et al.(1994) recently determined that APOE accounts for approximately 50% of the total genetic effect, suggesting that additional genes are important.

We have undertaken a genomic screen using a subset of 127 late onset AD cases in order to search for additional AD genes. We are using a 10cM sieve on 21 late onset families, with 3 or more living AD affecteds. Given the uncertainties in mode of inheritance, reliance on a single analytical method [such as lod score analysis] could result in a missed linkage. Thus, we are applying both model-free (Sibpair, APM) as well as likelihood methods to the problem.

Promising regions for follow-up are based on a nominal p-value of 0.05 for the model-free methods or a lod score > 1.0. In order for a result to be followed up it must fit on 2 of the above criteria. Adjacent markers fulfilling a single criterion also warrant follow-up. To date we have screened 112 markers on chromosomes 1, 2, 3,4, 5, 8, 11, 16, 18, 21, and 22. Two regions have satisfied our criteria for follow-up and these regions are currently being investigated further as we simultaneously continue with the genomic screen. Updated results of all analyses will be presented.

34

Toward Automating Likelihood Maximization

LD Atwood[1], AF Wilson[2], WC Buckalew[3]

[1]Southwest Foundation for Biomedical Research, San Antonio, [2]Louisiana State University Medical Center, New Orleans, [3]California Polytechnic State University, San Luis Obispo.

One of the practical difficulties in segregation analysis is adequately searching the likelihood surface. The likelihood surface must be thoroughly explored to ensure that the global maximum, as opposed to a local maximum, has been found. Traditionally, this has been accomplished using multiple maximizations, each with a different set of initial values.

The ongoing advances in the speed of computers now makes it possible to consider a strategy of random initial values in a large number of maximizations. We have modified the SAGE REGC program to generate uniform random initial values in the domains of the parameters. Initial values for the means are generated from the bounds on the data. All covariates are started at zero. These default values may be overridden.

We tested a 15 parameter codominant model with arbitrary transmission probabilities, arbitrary familial correlations, and three covariates on a data set with 674 subjects in 16 families. On a SPARCstation 20 this model averages 410 seconds per run. On a DECstation 3000/600 the same model averages 111 seconds per run. This approach provides a thorough exploration of the likelihood surface and relieves the analyst of a large amount of tedious work.

Supported by NIH grants HL28972, HL45522 and RR09919.

35

Estimation of the additive variance : Half-sibs children fathered by artificial insemination by donor as a model. Capron, C.,[1] Eliasziw, M.,[2] Czyglik, F.,[3] Jouannet, P.,[3] & Duyme, M.[1]

(1) EP12 - CNRS, Montpellier and Laboratory of Epidemiology Genetic, University Paris VII, France; (2) Department of Epidemiology and Biostatistics, University of Western Ontario, Canada; (3) CECOS, Port Royal Hospital, Paris, France.

The use of a design involving paternal half-sibs fathered by the same donor by artificial insemination to different women offers an opportunity to estimate the genetic components of some complex traits. Here, the model is tested on birthweight. The genetic model assumes mothers have nothing in common with each

other, except the general common environment during the pregnancy (environment between families), and there is no assorted mating. The design enable us to eliminate 1/ prenatal environment intra-family and 2/ prenatal maternal effects, e.g. cytoplasmic inheritance, uterine environment. Thus the resulting degree of resemblance among half-sibs is due to the additive effects of genes only and can be estimated by the intraclass correlation. Data from 154 donors yielding 537 half-sibs born from 537 different women have been collected. The 154 donor-groups differ in the birth rank, the sex and size of sibships. The birthweight data were converted to standardized scores using a multiple regression analysis to adjust for the duration of gestation, rank order, and sex of child. The adjusted ANOVA estimate of the degree of resemblance was 0.069 (p-value=0.04). Comparisons with results arising from other types of designs that estimate the additive variance for the birthweight, are discussed.

36

Genomic screening in multiple sclerosis. JL Haines1, DE Goodkin2, J Rimmler4, R Lincoln2, A Bazyk1, J Oksenberg2, SL Hauser2, MA Pericak-Vance4. 1). Massachusetts General Hospital, Boston; 2). University of California, San Francisco; 4). Duke University Medical Center, Durham, NC.

Application of positional cloning to diseases with a complex etiology is fraught with problems including undefined modes of inheritance, heterogeneity, and epistasis. Although microsatellite markers make genotyping the genome a straightforward task, genetic analysis of this data is not as easy. For datasets with homogeneous pedigree structures (such as only sibpairs or only large, extended pedigrees), the use of a single analytical method may be viable. However, for diseases such as multiple sclerosis (MS) with mixed pedigree structures, no single analytical method is available to efficiently and accurately analyze these data. We have developed a multi-stage, multi-analytical genomic screening strategy which uses a combination of non-parametric approaches (Affected Pedigree Member (APM) linkage analysis and robust sib pair analysis (SP)), and the parametric lod score approach (using four different genetic models). To warrant follow-up, a marker must have two or more of: a nominal P value of 0.05 or less on the non-parametric tests, or a lod score of 1.0 or greater for any model. Two adjacent markers each fulfilling one criterion are also considered for follow-up. These criteria were determined both by simulation studies and our empirical experience in screening a large number of other disorders.

We applied this approach to multiple sclerosis (MS), a complex neurological disorder with a strong but ill-defined genetic component. Analysis of the first 139 markers (on 14 chromosomes) from our screen of 55 multiplex families found 9 markers which met the SP criteria, 18 markers which met the APM criteria, and 9 markers which met the lod score criteria. Six regions (on chromosomes 2, 4, 5, 6, 7, and 19) met our overall criteria. However, no single method identified all of these regions, and each method identified regions that no other region supported. Only one region was identified by all three methods (chromosome 19). This suggests each method is sensitive to various (unknown) influences.

38

THE INHERITANCE OF FACTORS ASSOCIATED WITH JOINT MOBILITY.

M.M. Moura[1], M.F. Feitosa[1], J. Laredo Filho[2] and H. Krieger[3].
[1] Departamento de Genética, Instituto Oswaldo Cruz, Rio de Janeiro; [2] Departamento de Ortopedia, Escola Paulista de Medicina, São Paulo, [3] Departamento de Parasitologia, Universidade de São Paulo, São Paulo, Brazil.

The strong association between joint mobility with some Human diseases and syndromes support the search for the background causes of the variability of this phenotype. On a sample of 9127 individuals belonging to 1806 migrant nuclear families from Northeastern Brazil, information on the inter-phalangial mobility was obtained: a) The grades of flexion of both the right and left thumbs and b) The angle (in degrees) formed by the distant and proximal phalanx of the thumb. The first principal component of these variables was estimated and called "extensibility." A negative association of extensibility and age as well as with inbreeding were detected. Complex segregation analysis was applied to extensibility and both a polygenic mechanism and another transmissible component were detected. The Mendelian inheritance was rejected (χ^2_3= 39.5, P < 0.001), while a model with multifactorial inheritance together with a factor that is inherited with a transmission probability different from 1/2 (τ = 0.62) is not rejected (χ^2_2= 2.13, P = 0.345). These findings were supported by path analysis, which showed an important biologic inheritance (h^2 = 0.675 ± 0.024, χ^2_1= 160.91, P < 0.001) and the existence of a small but significant cultural component (c^2 = 0.0025, χ^2_1= 5.15, P = 0.023). The observed "inbreeding" effect therefore, could not be attributed to a genetic mechanism and probably is the effect of concomitant environmental variability.
(Supported by CNPq and FAPESP).

39

DESIGN EFFECTS FOR FAMILIAL CORRELATION IN REGRESSION MODELS

SB Bull, CMT Greenwood, NH Chapman, GA Darlington, Dept of Preventive Medicine and Biostatistics, University of Toronto

In genetic epidemiologic studies, the relationship between disease traits and explanatory covariates, such as candidate genes, is often evaluated with generalized linear models. Standard models assume independent observations, which is unlikely to be true for family data, and the usual standard errors may be too large or too small, depending on the values of the covariate. Generalized estimating equations models provide one method to adjust for such correlation. The effect of familial correlation can be measured by an adjustment factor, sometimes referred to as a design effect, which is the ratio of the adjusted variance to the variance calculated as if the family members were independent.

In this presentation, we examine typical design effects for genetic epidemiology study designs using an analytic approximation given by previous authors, and illustrate how the design effect is influenced by the intra-familial distribution of covariate values such as would be expected for genotype. These approximations can also be applied in other genetic analyses that involve dependent data, such as relative-pair linkage studies. Design effect values may also be useful at the design stage in correcting sample size requirements.

40

EXPECTATIONS FOR THE MULTIALLELIC TRANSMISSION-DISEQUILIBRIUM TEST (TDT).

David L Duffy. Johns Hopkins University School of Public Health, Baltimore MD; Queensland Institute of Medical Research, Brisbane, Australia.

A number of authors have now presented methods for the extension of the TDT (Spielman et al. Am J Hum Genet 1993; 52:506) to multiallele markers. However, the justification for these approaches is often heuristic, and a formal derivation is yet to be presented. Here I present the expectations for the multiallele TDT under a generalised single major locus model for the linked/associated trait, both in its usual form, and as a "genotypic" test that compares the proband genotype to the "non-transmitted genotype" produced by the union of both parents' nontransmitted alleles. The "genotypic" TDT also addresses the question of how to combine parental contributions from the same nuclear family. Although these formulae are complex, the test statistic for no association and no linkage is a straightforward test for symmetry in the NxN table of transmitted and nontransmitted alleles or genotypes.

41

Autistic behaviors in young girls with fragile X syndrome.

D.V. Dawson and A. M. Lachiewicz.
Case Western Reserve University School of Medicine, Cleveland, Ohio, and Duke University Medical Center, Durham, North Carolina, USA.

Mothers of 59 girls with fragile X [fr(X)] syndrome and 58 controls were asked to complete the Autism Behavior Checklist (ASIEP Education Co., Portland, OR) to help assess whether girls with fr(X) syndrome display more autistic behaviors than other girls with developmental disabilities. Disabilities of the controls included PKU, Down syndrome, and tuberous sclerosis; girls referred for fr(X) testing with negative results were also included. There were no significant differences between the groups in age or IQ. Mean age of fr(X)+ girls was 7.5 years vs. 7.1 years in controls; median IQ scores were 77 vs. 71. The distribution of total autism scores was significantly higher ($p<.0001$) among fr(X)+ girls (median 32 vs. 12 in controls. Fourteen fr(X)+ girls exhibited scores $\geq$ 67, a value which has been cited as a probable cutpoint for autism, whereas none of the controls did. Low scores were seen in both groups, but more frequently in controls. Fragile X girls also had significantly higher scores on the five subscales: Sensory, Relating, Body/Object Use, Social/Self-Help, and, to a lesser extent, Language. Significance persisted after adjustment for IQ and age group. These findings suggest that about 25% of girls affected with fragile X may have autistic behaviors as a feature of the disability, but that these behaviors do not occur in all girls.

42

Maternal risk factors for major associated anomalies in infants with Down syndrome. CP Torfs and RE Christianson, California Birth Defects Monitoring Program, Emeryville, CA 94608-1811.

Infants with Down syndrome (DS) have specific additional birth defects such as atrioventricular canal (AVC) or duodenal atresia (DA) in higher proportion than the remainder of the birth population. Khoury and Erickson (1992) suggested that maternal demographic factors and pregnancy exposures might interact with an aneuploid genotype to modify its susceptibility to those defects. To test this hypothesis, we used: a) a 1991 to 1993 case control study of 693 DS live births whose mothers were interviewed; b) the 1983 to 1992 CBDMP data base that provided us with rates of defects in over 2,200,000 births, excluding infants with DS. We compared maternal demographic risk factors for the same defect in both data sets and evaluated maternal exposure factors in the case control study. Among the 693 DS live births, 52% had a cardiac defect (15.6% AVC, 17.6% ASD without AVC; 16.2% VSD without AVC). Smoking was associated with cardiac defects (OR=2.04; CI=1.25-3.30), AVC (OR=2.07;CI=1.2-3.58), ASD (OR=1.67;CI=0.92-3.0), and with DA (OR=2.01; NS). In contrast to Khoury's results, none of the mothers of infants with DA (N=29) had fever during the first trimester. Although Blacks had a lower prevalence of AVC than Whites, Hispanics, or Asians among infants without DS, among infants with DS Blacks had a higher prevalence (24%) of AVC than other races (14% to 20%). These results show that maternal and environmental factors can modify risks for defects associated with DS.

*Khoury MJ & Erickson JD, Am J Med Genet 43:1016

43

Identification of linkage phase.

Axenovich T.I.

Institute of Cytology and Genetics, Novosibirsk, 630090 Russia

The most popular way to get linkage phase data, is analysis of the parental disease statuses and marker genotypes. Unambiguous identification of

the linkage phase is usually demonstrated on dominant diseases. For more complicated cases that require introduction of incomplete penetrances, linkage phase identification is of special importance as less information is provided about linkage. I demonstrate that i) the parents with the same disease status or marker genotype carry no information about the linkage phase of the diheterozygous offspring; ii) the linkage phase can be inferred unambiguously only under recessive or dominant control of disease. In any other situation, only the probability of linkage phase is inferred. I denote the accuracy with which the linkage phase can be predicted by the index x which is equal to the magnitude of the relative difference between the probabilities of coupling phase and repulsion phase. The value of x has been shown to be the same in all informative crosses and to be dependent on genetic model of the disease. Besides, x is the same for the di- and polyallelic marker genes. I have evaluated x and assessed the frequencies of informative crosses under different genetic models of the disease.

44

Pedigree data management and analysis using the Genetic Analysis Package

W.J. Gauderman, C.L. Faucett, J. Buckley, J.L. Morrison, and D.C. Thomas.
University of Southern California, Los Angeles

The Genetic Analyis Package (GAP) is a comprehensive computer program for the management and analysis of pedigree data. The data-management component provides a convenient tool for importing, viewing, editing, and printing pedigrees and associated data. The analysis component includes the following features:

- Continuous, binary, or censored age-at-onset phenotypes
- Combined segregation and linkage analysis
- Inclusion of measured covariates and GxE interactions
- Major-gene, polygene, or mixed models
- Correction for ascertainment
- Analysis of complex pedigree structures, including loops
- Estimation by Gibbs sampling for all models
- Estimation by maximum likelihood, where possible

The data management and analysis programs of GAP are fully integrated in a user-friendly, menu-driven environment. GAP is currently under development but will soon be available, for both PC- and Sun-based operating systems. We will demonstrate the essential features of GAP, concentrating on the models, methods, and assumptions used in the analysis component.

45

A century of mortality in three large families with Huntington's disease

E.T.M. Hille[1], R.A.C. Roos[2], M. Vegter-van der Vlis[3], F.R. Rosendaal[1], J.P. Vandenbroucke[1], Dept. of Clinical Epidemiology[1], Neurology[2] and Clinical Genetics[3], Leiden University Hospital, The Netherlands

Huntington's disease (HD) is an autosomal dominant disease characterized by chorea, dementia, personality changes and emaciation. In this study we compared the mortality in members of three large HD families (observed) with that of the general Dutch population (expected) adjusted for age, sex and calendar period. In addition the effect of paternal or maternal transmission of the HD gene on mortality ratios was determined. The ratio of observed to expected is the standardized mortality ratio (SMR), a measure of relative risk. We used Mendelian reasoning to identify all members of the families with a 50% or 100% prior probability of carrying the HD gene in the past. After removal of other relatives 272 individuals contributed to the analysis. Follow-up extended from the date of birth of affected offspring to the end-of-study date.
On a total of 7000 person-years 116 deaths were counted, whereas the expected number of deaths was 73.4 (SMR 1.6, 95% CI 1.3-1.9). For paternal inheritance the SMR was 2.1 (95% CI 1.6-2.5) and for maternal inheritance 1.4 (95% CI 1.0-1.9). In a Poisson regresssion model this difference was significant (rate ratio 1.7; 95% CI 1.1-2.6). The excess mortality was strongest in the age group 40-59 years (SMR 2.8, 95% CI 2.0-3.6). Therefore, we studied the secular evolution of mortality in this age group. There was a continuous decline of mortality in the general Dutch population, but among the HD families this decline was absent. Because of this the rate ratios increased from 1.5 (95% CI 0.9-2.3) in the period 1870-1909 to 5.4 (95% CI 3.3-8.1) in the period 1975-1994. We conclude that the disadvantage of carriership of the HD gene in a family has greatly increased over time, relative to the general population.

46

Linkage analysis of complex psychiatric disorders: a new approach using Recombinant QTL Introgression (RQI) animal model systems.

Csaba Vadasz and Arthur Fleischer, Laboratory of Neurobehavioral Genetics, Nathan Kline Institute for Psychiatric Research, New York University Medical Center, 140 Old Orangeburg Road, Orangeburg, NY 10962

The turning point in the linkage studies of psychiatric disorders is coming with the recognition of the nonlinear interaction of multiple genetic and environmental factors influencing the development of most behavioral disorders (Cloninger, C.R., 1994, Am. J. Med. Genet., 54, 83-92). Appreciation of the formidable difficulties in the mapping of psychiatric disorders and recognition of the evolutionary conservation of genes are directing attention to animal behavioral gene-mapping studies. Here, we propose a novel approach, which is based on the transfer of trait-specific QTLs onto a homogeneous background and random distribution of the QTLs in RQI strains (Cs. Vadasz, I. Sziraki, L. R. Murthy, M. Sasvari-Szekely, P. Kabai, I. Laszlovszky, A. Fleischer, B. Juhasz, and R. Zahorchak, 1994, Mammalian Genome, 5, 735-737). Results of the construction of a 10 cM map of the first two RQI strains of our dopamine system-specific B6.CM_5 set suggest that the RQI approach would provide the most sensitive tool for location of QTLs. In both RQI strains, chromosomes #2 and #18 carried identical alleles. Testing of the hypothesis that the identified common segments carry dopamine system-specific QTLs is in progress. Identification of dopamine system-specific QTLs can be instrumental in mapping human QTLs for psychiatric disorders characterized by imbalance in the mesotelencephalic dopamine system.

47

TWIN_GEN: A SAS macro for the univariate genetic analysis of twin data. J. Goldberg[1,2], K. Bukowski[1], V. Ramakrishnan[1,2], J. Meyer[3], M.E. Vitek[1], M. Biondic[1], W.G. Henderson[1].
[1]Vietnam Era Twin Registry, Hines, IL; [2]UIC-SPH Epid-Biostat, Chicago, IL; [3]Human Genetics, MCV, Richmond, VA.

TWIN_GEN is an interactive SAS macro developed for the univariate genetic analysis of twin data. The objective of this analysis is to estimate and test for heritability and shared common environmental effects using within-pair measures of twin similarity. Input to this program must contain twin pairs (ie, each observation represents a single twin pair). The program permits analysis of dichotomous, ordinal or continuous twin characteristics. For dichotomous data the analysis uses the test statistics proposed by Ramakrishnan et al. (1992, Genet. Epi. 9,273-287); for ordinal (>2 categories) and continuous data the program uses an adaptation of the regression method proposed by DeFries and Fulker (1985, Behav. Genet. 15,467-473). With dichotomous or ordinal data the user has the option of using raw data files or directly inputing contingency tables. **TWIN_GEN** can be run via a double-click on a macro icon defined for SAS PC. This executes a SUBMIT command while in SAS for Windows (vs. 6.08). In order to keep abreast of the program status, formatted windows are displayed throughout the program run. This provides a friendly, user-driven application. Included in this macro are prompts for input/output filenames, online data validation, and various exit options. Statistics are retained and used to create a customized output table per each twin characteristic specified.

48

Familial Effects on Growth Rates of Very Low Birthweight (VLBW) Infants During the First 3 Years of Life. T. Yamashita, J. Huang, J. Baley and L. Singer (Department of Pediatrics, Case Western Reserve University, Cleveland, OH)

Twenty-nine twins (Ts) and an equal number of matched control pairs were analyzed to determine the familial correlation of growth rates in terms of weight and height changes during the first 3 years of life. Growth rates were calculated by fitting a regression model to the heights and weights of each individual who had 3 to 5 measurements at various ages. There were 8 bronchopulmonary dysplasia (BPD), 14 very low birthweight (VLBW) without BPD and 7 BPD-VLBW Ts; and 11 female (FF), 8 male (MM) and 10 FM pairs. Controls (C) were selected by matching a T with a non-twin by group (BPD/VLBW), socio-economic status, race, gender and mother's age. Pearson's correlation coefficients were used to identify association between T-T, T-C and C-C pairs. When the T-T correlation was significant ($p<0.01$) and others were not, we interpreted the familial correlation as significant. There were no birthweight, expected height and weight diferences between these pairs at 3 years. However, T-T correlations of the birthweights ($r=0.74$, $p<0.0001$), and the weight and height changes ($r=0.81$, 0.93, $p<0.0001$) were significant. Analyses within gender and group pairs showed similar relationships. Conclusion: Familial factors, whether genetic or environmental, appear to play an important role on growth rates among high risk (BPD/VLBW) infants. (Funded by: MCJ390592 & NIH-HL38193)

49

Heterogeneity of sarcoidosis risk among siblings and parents of sarcoidosis cases
B. Rybicki, M. Major, M. Maliarik, M. Simoff, J. Popovich, Jr. & M. Iannuzzi. Henry Ford Health System, Detroit MI.

The familial clustering of sarcoidosis cases and its higher prevalence and more severe clinical course in African Americans suggests an etiologic heterogeneity. To test for heterogeneity in familial risk of sarcoidosis, we studied 3,395 siblings and parents of 558 index cases diagnosed at Henry Ford Hospital between 1951-1994. Using the age- and sex-specific cumulative incidence of sarcoidosis in our sample, we found a statistically significant heterogeneity in familial risk of disease ($p=0.006$). To determine whether this was due to a greater risk of sarcoidosis in African Americans, who made up a majority (65%) of the sample, we recalculated disease probabilities using age-, sex- and race-specific disease cumulative incidence and found a similar high degree of heterogeneity in familial risk ($p=0.003$).
The two characteristics of index cases which had a positive association with high risk families, defined as having one or more affected family members ($n=69$), were African American race (Odds ratio (OR)=3.24; 95% confidence interval (CI)=1.71-6.14) and having an additional relative affected with sarcoidosis (OR=2.71; 95% CI=1.33-5.53).

We conclude that the heterogeneity of sarcoidosis risk in families of index cases found in this study is not supportive of multifactorial inheritance. Since African Americans make up a greater percentage of high risk families, they should be targeted for family studies which can further validate these findings by testing Mendelian hypotheses and for genetic linkage.

50

Evaluation of Population based prenatal diagnosis by a registry of congenital anomalies.
Stoll C, Dott B, Alembik Y, Roth MP.
Institut de Puériculture, Centre Hospitalo-Universitaire, 67000 Strasbourg, France.

Prenatal diagnosis performed by fetal karyotype and ultrasound scan is now a routine part of antenatal care in many countries. How many fetal anomalies are actually detected by these procedures ? We have used our registry of congenital malformations to answer this question. The population of malformed patients came from 186,466 consecutive pregnancies of known outcome from 1979 to 1992. In this population ultrasound scanning during pregnancy is routine

practice and fetal chromosomal examination is offered free of charge to women 38 years of age or more. During the study period 316 chromosomal anomalies were diagnosed including 207 Down syndrome patients. Termination of pregnancy was performed in 87 cases including 59 fetuses with Down syndrome.

The detection rate and the specificity of prenatal diagnosis by ultrasonographic examination were improved from 1979 to 1992. During the periods 1979-1988 and 1989-1992 prenatal diagnosis was performed in 20.1 % and 30.6 % of fetuses with non chromosomal anomalies, respectively. The detection rate for isolated malformations (fetuses with only one anomaly) and for multiple malformed children was 26.3 and 65.8, respectively from 1989 to 1992. This detection rate was variable for the different categories of malformations. A high detection rate was observed for anencephaly and urinary tract malformations and a low detection rate for limb reduction defect and for cleft lip.

51

Maximum likelihood estimation in linkage heterogeneity models including additional information via the EM-algorithm

Jeanine J. Houwing-Duistermaat[1]
Lodewijk A. Sandkuijl[2]
Hans C. van Houwelingen[1]
[1]Department of Medical Statistics
[2]Institute of Human Genetics, Leiden University

A set of 15 pedigrees with X-linked recessive retinitis pigmentosa will be analyzed. The penetrance of the disease is equal to one. For every pedigree a lodscore curve is known. These 15 curves show locus heterogeneity.

Some of the obligate female carriers show a tapetal reflex. The penetrance of tapetal reflex in the pedigrees also shows heterogeneity and may be associated with the type of locus of the disease gene. When such an association exists, the additional information can be used to get more evidence for locus heterogeneity.

A model based on the joint distribution of the position of the disease gene and the additional information will be presented. Here, it is used that only obligate female carriers are observed for tapetal reflex. It will be shown how to estimate the model parameters using the EM-algorithm and how to obtain confidence intervals for the model parameters and for the posterior probabilities via profile likelihood.

52

POPULATION-BASED CASE-CONTROL-FAMILY STUDIES OF THE GENETIC EPIDEMIOLOGY OF BREAST, COLORECTAL & PROSTATE CANCER

John L. Hopper
The University of Melbourne,Department of Public Health & Community Medicine
200 Berkeley Street, Carlton, Victoria, 3053

Typical cancer genetics studies use families in which there are numerous cases of one or more cancers across multiple generations. The strategy has led to the identification of family cancer syndromes,such as hereditary non-poplyposis colorectal cancer (HNPCC),and of the genes segregating within these large kindreds. From the epidemiological viewpoint however, it is difficult to extrapolate findings to the population,in part due to the usually ad hoc ascertainment of 'interesting' families. We have been conducting studies of breast,colorectal and prostate cancer in which families are identified either through cases on the population-based Victorian and New South Wales Cancer Registries,through their spouses or partners, or through population based controls.Questionnaires are used to measure epidemiological risk factors,and DNA is collected. Strict rules define the cases (eg breast cancer diagnosed before age 40,living in Melbourne,and the relatives to be initially studied (eg first degree relatives, grandparents and aunts). Sampling of relatives is extended to those families in which cancers are occurring by the Cannings-Thompson Sequential Ascertainment Scheme, to the population by a simple adjustment This design makes it possible to measure the population burden associated with cancer genes, and to study concurrently the roles of both genetic and environmental risk factors on cancer susceptiblity. Such information is necessary for assessing the costs and benefits of potential genetic screening programs.

53

International Molecular Epidemiology Task Force (IMETAF): Biotechnology and Public Health Training in Developed and Developing Countries. J.S. Dorman (University of Pittsburgh, Pittsburgh, PA USA), C. Gorodezky (INDRE, México City, México), A. Libman (University of Rosario, Rosario, Argentina)

Molecular epidemiology is a new field which has emerged from the integration of advanced biotechnology into epidemiologic research. Its objectives include the evaluation of host/environmental interactions in disease, and the development of prevention approaches using molecular diagnostics in public health settings. To assist nations acquire expertise in molecular epidemiology, IMETAF was established in 1993. Its mission is to facilitate the development and implementation of programs in molecular epidemiology in all regions of the world, and to promote advanced biotechnology transfer for scientific research and its integration into medicine and public health for disease prevention. Two countries in Latin America, México and Argentina, have formally begun national molecular epidemiology programs. Although the specific activities of each program varies, they share common international objectives including: 1) developing multi-disciplinary collaborative networks with government support, 2) assessing major health problems amenable to molecular epidemiology, 3) conducting a situational analysis of the state-of-the-art in epidemiology and molecular biology, 4) implementing workshops and training programs in molecular epidemiology, and 5) establishing molecular epidemiology projects for disease prevention and control. IMETAF is currently expanding to include countries in Asia and the

Eastern Mediterranean Region. IMETAF will also assure that molecular epidemiology programs complement current national health policy and contribute to WHO's objective of "health for all by the year 2000".

54

The effect of reproductive risk factors on the age of onset of disease for BRCA1-carriers. Jenny Chang-Claude, Heiko Becher, Ute Hamann. German Cancer Research Center, Heidelberg, Germany

Introduction: Female carriers of the BRCA1 gene have a very high risk of developing breast and/or ovarian cancer during their lifetime. There is, however, little knowledge to what extent non-genetic risk factors such as age at menarche, age at first birth, abortion, alter the age of onset of the disease.
Study design: We conducted a family study of breast cancer, recruiting families with at least three breast cancer cases diagnosed under the age of 60. Blood samples were collected from all relevant family members for molecular genetic analysis. Information on known and suspected risk factors was collected from all female members using a self-administered questionnaire. All diagnoses of cancer were verified through pathology reports and paraffin blocks were obtained when available. 43 families were studied for linkage to the susceptibility locus, BRCA1.
Statistical methods: Multipoint LOD scores using flanking markers of BRCA1 were used to identify families showing probable linkage. Probable BRCA1-gene carriers were identified through the haplotypes which segregated with the disease in the family. We performed a survival analysis for probable gene carriers by calculating logrank tests and fitting proportional hazards models in order to identify risk factor groups.
Results: 59 gene carriers were identified of whom 42 were affected. Unaffected gene carriers were mostly young (median age 31 years) and were treated as censored at current age for the analysis. Median age at onset (breast cancer or ovarian cancer) for affected gene carriers was 42.5 years. As risk factors we considered age at menarche, age at first pregnancy and abortions. Compared to women with ages below 25 at first birth, women without pregnancies or with later ages at first birth were found to show a significantly different survival curve (logrank-test $p = 0.01$) and a significantly increased risk in the proportional hazards analysis. Age at menarche and abortions did not show a significant effect.
Conclusion: Our data provide some evidence that reproductive risk factors for breast cancer can have an effect on age at onset for BRCA1-gene carriers. However, considering that our analyses are based on limited numbers, these results warrant further clarification.

55

GENETIC ANTICIPATION IN FAMILIAL ADENOMATOUS POLYPOSIS (FAP)

S. Presciuttini, P. Sala, C. Rossetti, and L. Bertario.

Italian polyposis registry, Milan, Italy

Age of symptom onset is extremely variable in FAP. This is a Mendelian condition highly predisposing to colorectal cancer, caused by inactivating mutations in the *APC* gene. In our registry, the youngest symptomatic patient was a 5-year-old boy, and the oldest was a 79-year-old woman. Part of this heterogeneity is explained by pedigree factors, the particular mutation in the *APC* gene being of special interest, but most of the total variation is observed within families. When we considered this within-group component in a sample of 583 patients subdivided by family, we found that a very simple explaining factor was the patient generation index. For example, in three large pedigrees with identified mutation at codon 1309, the overall means in generations 1, 2 and 3 were 54.5 years (n=4), 30.7 years (n=13) and 15.5 (n=11), respectively, with a mean anticipation in subsequent generations of 19.5 years. We considered, for each individual of the total sample, the age of symptom onset against the year of birth, and measured the anticipation in terms of slope of regression line. The computed value was -.51. Considering a generation time of 30 years, for any random parent-child pair we expect the child to develop polyposis about 15 years earlier than his or her affected parent. Although we cannot exclude that ascertainment bias may contribute to this phenomenon, we suggest that genetic anticipation should be further investigated in FAP.

56

The Relationship Between Smoking Exposure and p53 Overexpression in Colorectal Cancer. A.N. Freedman, A.M. Michalek, J.R. Marshall, C.J. Mettlin, N.J. Petrelli, J.E. Asirwatham, J.D. Black, S. Satchidanand, Z.F. Zhang (Roswell Park Cancer Institute, Buffalo, New York 14263)

Background: Although epidemiologic studies of the relationship between cigarette smoking and colorectal cancer risk has been equivocal, a positive association is consistently found for colorectal adenoma development. Therefore, tumors acquiring gene alterations late in the adenoma-carcinoma (such as p53) may be unrelated to smoking exposure and may confound studies of smoking and colorectal cancer risk. **Purpose:** This study was designed to determine whether p53 protein overexpression is associated with cigarette exposure in colorectal cancer. **Methods:** 163 colorectal cancer cases and 326 healthy controls provided information concerning their history of cigarette smoking. All patients' tumors were analyzed immunohistochemically for p53 overexpression using an avidin-biotin immunoperoxidase procedure and polyclonal anti-p53 antibody CM1. **Results:** Comparison of colorectal cases versus controls revealed elevated but nonsignificant increases in risk for ex-smokers (OR=1.51, 95% CI 0.96-2.36) and current smokers (OR=1.22 95% CI 0.70-2.14) when compared to nonsmokers. A significant dose response relationship was found for total pack years of smoking (Trend test: $p = 0.05$). When p53+ cases were compared to controls, no associations were observed for smoking status or total pack years of smoking. However, when p53- cases were compared to controls, an elevated risk was found for ex-smokers (OR=2.07, 95% CI 1.15-3.74) and current smokers (OR=1.94, 95% CI 0.97-3.85). A significant trend of increasing risk was also found for total pack years of smoking (Trend test: $p = 0.03$).
Conclusion: Colorectal tumors developing through p53+ dependent pathways were not associated with smoking exposure. A significant increase in risk was observed for the p53- independent pathway with a patient's history of smoking.
Implications: p53 molecular subtypes are associated with smoking exposure in colorectal cancer.

57

Linkage studies at BRCA1 locus in 52 French breast and/or ovarian cancer families: implications for genetic counseling

Laurent Essioux, Dominique Stoppa-Lyonnet, Catherine Girodet, Michel Longy, Christine Maugard-Louboutin, Fabienne Kerangueven, Daniel Birnbaum, François Eisinger, Didier Lanõe, Claude Toulouse, Isabelle Holstein, Catherine Bonaïti-pellié, Yves-Jean Bignon, Hagay Sobol, and le Groupe Génétique et Cancer de la Fédération Nationale des Centres de Lutte Contre le Cancer.

Linkage analyses at BRCA1 locus in 52 French families with breast and/or ovarian cancers were performed to evaluate the information provided by BRCA1 markers for genetic counseling, and to propose recommandations for genetic testing of this predisposition. The sample included 35 breast-only families, 15 breast-ovary families and 2 ovary-only families containing at least three affected cases, first or second degree relatives. Risks of being a mutation carrier were calculated for unaffected first degree relatives of cases using three markers flanking the BRCA1 region: THRA1, D17S800 and D17S579. Genetic heterogeneity of the hereditary predisposition was taken into account in the analyses.
Before using information on BRCA1 markers, an evaluation of the probability that an affected individual carries a mutation suggested that in six breast-only families, the familial aggregation of cases was probably due to chance. Ten families had a posterior probability of linkage to BRCA1 higher than 0.8, the corresponding risks ranging from 0.02 to 0.97. In 7 of these families, risk values greater than 0.9 or less than 0.1 were obtained, allowing a determination of the genetic status of at risk individuals with a sufficient degree of certainty. Only one of these families belongs to the breast-only group, displaying seven cases of breast cancer, and one to the ovary-only group. The five other families display at least 2 breast cancer cases in addition to the ovarian cancer and at least 3 affected members typed (or inferred).
Our study illustrates the interests and limits of the indirect determination of carrier status of individuals in the breast/ovary cancer predisposition. This approach will be improved when the location of BRCA2 is precised. In genetic counseling, we recommand to undertake a family investigation when the pedigree displays some necessary characteristics, in terms of number of affected cases, relationship between them and availability of family members, which are described.

58

Lung cancer risk and genetic polymorphism of GSTM1 among African-Americans and Caucasians. London SJ, Daly AK, , Leathart J, Navidi WC, Carpenter CL, Idle JR. USC School of Medicine, Los Angeles CA 90033 and Univ. Of Newcastle, U.K.

A common polymorphism of Glutathione S-transferase class μ (GSTM1) leading to variation in the ability to detoxify certain carcinogens has been proposed as a risk factor for lung cancer. However, published data are inconsistent. We examined the association between the GSTM1 null genotype and lung cancer risk in a study of 342 cases and 716 population controls of African-American and Caucasian ethnicity in Los Angeles County. For all lung cancer combined, the GSTM1 null genotype was associated with an odds ratio (OR) of 1.18 (95% confidence interval (CI) 0.87-1.60). The odds ratio did not vary materially between African-Americans (OR = 1.13) and Caucasians (OR = 1.25). The association was strongest for squamous cell carcinoma (OR = 1.45, 95% CI 0.88-2.39). The odds ratio differed significantly according to lifetime smoking history (p=0.038) with an association limited to smokers of less than 40 pack-years (OR = 1.49, 95% CI 0.97-2.29). In conclusion, our data do not support a major association between GSTM1 null genotype and lung cancer overall in either African-Americans or Caucasians. Modest associations may be present within subsets of squamous cell carcinoma and lighter smokers although limited power within these categories preclude stable estimates. (Supported by TRDRP 3RT-0403).

59

Is there evidence for genetic factors controlling mammographic breast density?

Vachon C., Kuni CC, King RA, Rich SS, Anderson VE, McGovern PG, Sellers TA. University of Minnesota

Breast density has been shown in previous studies to be positively associated with breast cancer risk. Few studies, however, have addressed the familial aggregation of breast density. We estimated familial correlations of breast density for mother/daughter pairs and sister pairs in a cohort of 544 Minnesota breast cancer families initially ascertained between 1944 and 1952. Original mammogram films were read by a single radiologist who visually estimated percent density. Familial correlations were derived using the FCOR program in SAGE. The equal weighting to pairs method presented here, is based on 43 mother/daughter pairs and 150 sister-pairs with breast density readings. To assess statistical significance, we used Fisher's Z-transformation for the mother/daughter interclass correlations and an asymptotic variance test for the sisters' intraclass correlations. Simple and age-adjusted correlations were 0.11 and 0.13 respectively, for mother-daughter pairs; the corresponding correlations for sister pairs were 0.25 and 0.16. Only the simple sib-pair correlation of 0.25 was statistically significant($p<.05$). Subsequent analyses were performed to adjust singly for several risk factors, in addition to age, that our previous analyses suggested were correlated with density. These included reproductive variables (age at first birth, age at last birth, age at menarche, age at menopause, pregnancy history, number of pregnancies, menopausal status, hormone and oral contraceptive use), anthropometric variables(body mass index, waist-to-hip ratio) and lifestyle variables(alcohol, smoking, exercise). There was a significant mother-daughter correlation after adjustment for age and the number of pregnancies($r=.33$)($p<.05$). For the sister pairs, the adjustment for age and alcohol use resulted in a statistically significant correlation($r=.21$)($p<.05$). These results, albeit based on small numbers, provide little evidence of a strong familial component of breast density.

61

Family History and Risk of Developing Bilateral Breast Cancer among Young Women.

JL Bernstein[1], WD Thompson[2], N Risch[3], TR Holford[4], D Bruce[4]. [1] Mount Sinai Medical Center, New York, NY, USA. [2] University of Southern Maine, Portland, ME, USA. [3] Stanford University, Stanford, CA, USA. [4] Yale University, New Haven, CT, USA.

The purpose of this prospective cohort study was to evaluate the risk of developing second primary breast cancer among women under age 55 who were diagnosed with a first primary breast cancer between 1980 and 1982. Cases

used for this study were part of a multi-center, population-based case-control study and were followed through the end of 1989 for the occurrence of second primary breast cancer. From the cohort of 4,678 women, 330 developed a second primary breast cancer in the contralateral breast. Cox proportional hazards modelling techniques were used to model time to onset of the second primary, while adjusting for multiple predictors. The risk of developing a second primary breast cancer according to family history of cancer was dependent on the age of diagnosis of the first primary; compared with women diagnosed with their first primary after age 45, women diagnosed early had a substantially elevated risk. Specifically, women diagnosed at age 45 or younger had an increased risk of developing bilateral breast cancer if they reported having a first degree relative with: ovarian cancer (RR=3.35, 95% CI 1.22-9.18); endometrial cancer (RR=1.97, 95% CI 1.01-3.87); and breast cancer (sister: RR=4.13, 95% CI 2.26-7.55: mother (diagnosed at age 45 or younger): RR=3.55, 95% CI 1.41-8.93; sister and mother (diagnosed early): RR=18.61, 95% CI 4.59-75.60). For women diagnosed after age 45, the RR's were smaller and not significantly greater than 1.00. These findings from a prospective study suggest that women with inherited susceptibility to breast and ovarian cancer are at greatest risk for developing bilateral disease.

62

Lack of Evidence for Familial Aggregation of Non-Hematological Neoplasms in Relatives of Patients With Hematological Neoplastic Disorders. O. Shpilberg, M. Modan, B. Modan, A. Chetrit, B. Ramot. Institutes of Hematology and Clinical Epidemiology, Sheba Medical Center, Tel-Hashomer 52621, Israel

Familial aggregation of hematological neoplasms (HN) and non-hematological neoplastic disorders (NHND) were compared in 189 families of patients with HN with 2 control groups of 36 families of patients with benign hematological disorders and 33 families of patients with diabetes mellitus (DM). A self-administered questionnaire was used requesting a full list of 1st and 2nd degree relatives, their vital status, current age or age at death, and their chronic diseases. Previously, we have reported a significantly increased odds ratio for HN among 4061 relatives of HN patients versus the 1463 relatives of the controls (odds ratio 3.61, 95% confidence interval 1.44-9.07, $p<0.01$). [Brit J Haematol 87;75-80, 1994]. However, the current analysis shows no evidence for significantly increased tendency for developing NHND in relatives of patients with HN in comparison to those of the control group (61/1000 relatives versus 43/1000 relatives, respectively; $p=0.5$). In multivariate analysis, an odds ratio of 0.88 (95% confidence interval 0.61-1.27) was calculated when the HN group was compared to the benign hematological disorders group, and 0.84 (95% confidence interval 0.54-1.29) when the HN group was compared to the DM group. The most prevalent NHND among HN relatives were gastrointestinal neoplasms (12.8/1000), breast cancer (6.7/1000), and pulmonary neoplasms (5.7/1000). These figures were lower, yet not statistically significant, in the control groups. We conclude that increased aggregation of malignant disorders among relatives of patients with HN is unique to the hematopoietic system and not to other types of neoplasia, and might be due to genetic predisposition to HN in these families.

63

A Validation Study of Family History of Breast Cancer from a Population-Based Cancer Registry. Hoda Anton-Culver, Tom Kurosaki, Thomas H. Taylor, Epidemiology Division, University of California Irvine

A major risk factor for breast cancer is family history of the disease in first-degree relatives. In etiologic investigations, family history of cancer is typically ascertained through personal interview. Although family history information is often readily available to cancer registries through medical charts, the data are of unknown validity. This study evaluates the validity of family history information on breast cancer in mothers and sisters of breast cancer probands from the cancer registry (CR) compared to personal interviews (PI) of 316 consecutive cases of breast cancer. Breast cancer is seen in mothers of 14% of probands by CR compared to 12% by PI. Further, 16% of probands have a sister with breast cancer using CR compared to 13% by PI. The higher percentages from CR compared to PI may be due to non-recording of negative family history in the chart. PI data show only 6% of those with no family history recorded in the CR are positive for mothers with breast cancer, and a similar figure is seen for sisters. Using the PI as the standard, the sensitivity of CR to detect a breast cancer affected mother is 90% and the specificity is 99%. Similarly for breast cancer affected sisters the sensitivity is 91% and the specificity is 98%. For proband-mother and proband-sister breast cancer, CR and PI data yield similar age-specific prevalence rates. The age groups considered are 30-49, 50-64, 65-79, and 80+ years. For proband-mother pairs, the age-specific rates from CR are 15.1, 12.7, 10.6, and 4.6%, respectively; from PI these rates are 17.0, 9.6, 12.5, and 6.6%, respectively. For proband-sister pairs, the CR are 4.3, 6.5, 13.6, and 9.9%, respectively; from PI these rates are 2.4, 5.7, 14.5, and 7.4% respectively. In conclusion, the sensitivity and specificity of using CR data are high and estimates of age-specific prevalence of proband-mother and proband-sisters with breast cancer from CR and PI are sufficiently similar to warrant the use of CR family history data in studies of genetic epidemiology.

64

Segregation analysis of breast cancer: A comparison of type-dependent age-at-onset versus type-dependent susceptibility models

((Ciske D, Rich SS, King RA, Anderson VE, Bartow S, Vachon C, McGovern PG, Kushi LH, Zheng W, Sellers TA,.)) Division of Epidemiology, Univ. of Minnesota, Minneapolis, Minnesota

Most segregation analyses of breast cancer susceptibility have modelled the effect of the major gene on the age-at-onset distribution. In families linked to BRCA1 or BRCA2 however, there is wide variation in the age at onset among gene carriers. We performed a segregation analysis of breast cancer using models which parameterized the putative major gene effect in two ways: earlier age at onset, with a common level of susceptibility (Model I), and greater susceptibility, with a common mean age at onset (Model II). A total of 544 Minnesota families (n=12,732 members) were ascertained through breast cancer probands diagnosed between 1938 and 1952. Maximum likelihood segregation analyses were performed using the REGTL program in S.A.G.E. to fit five hypothetical modes of transmission and an unrestricted general hypothesis to the data. Twice the difference between the $\log_e$ likelihood for the data under the specified hypothesis (recessive, no major gene, etc.) and the $\log_e$ likelihood under the general hypothesis was compared to the χ^2 distribution to assess goodness-of-fit.

Under Model I, both Mendelian and non-Mendelian hypotheses were rejected. When Model II was used, the non-Mendelian hypotheses were rejected whereas all Mendelian hypotheses were not. Mendelian recessive inheritance of a common allele ($q_A = 0.11$) with a high penetrance (87%) provided a slightly better fit to the data than dominant or codominant inheritance. We then stratified the families into two subsets based on the age at diagnosis of the proband (≤ 55 years (n=262) versus > 55 (n=275)); there was no evidence of heterogeneity under either Model (I or II). According to Akaike's information criterion, the recessive and codominant hypotheses fit the data better under Model II, than under Model I. These data suggest that, while major genes greatly increase lifetime risk of breast cancer, non-genetic factors (diet, reproductive factors, etc.) may be more important determinants of age at onset.

65

Estimating population-specific disease risks for a given family history with an application for breast cancer. Heiko Becher, Jenny Chang-Claude. German Cancer Research Center, Heidelberg, Germany

Introduction: For many chronic diseases a family history is known to be a strong risk factor and it is of interest to estimate the disease risk given a particular family history. For the American white population tables have been published which give estimated age-specific cumulative breast cancer risk under various family history situations. As the risk for chronic disease is usually a mixture of exogenous and endogenous factors, such tables are not directly applicable to other populations.
Methods: We present a method to calculate appropriate tables for disease risks given a particular family history for arbitrary populations. We assume that such tables are available for a particular population. The method is based on two steps 1.

Decomposition of the tables with the family disease risk estimates for population Ω into probabilities for carrying the disease gene given a particular family history and for disease probabilities by gene carrier status. 2. Suitable composition of these probabilities to obtain risk estimates for a different population, Ω_2. We need the assumption that (i) there is (are) autosomal dominant disease gene(s) and that the probability of the disease given the disease gene is independent of the population and the family history and (ii) the frequency of the disease gene is independent of the population. Assumption (ii) is necessary because estimates of the gene frequency for different populations may not be available, however, direct information on population based gene frequencies, if available, can directly be incorporated into the method.
Example: For breast cancer, some 5% of all cases are due to inherited autosomal dominant disease genes, two of which have been identified (BRCA1) or localized (BRCA2). The majority of the breast cancer cases, however, are sporadic, thus accounting for large differences in the occurrence rates between populations. For example, Japan has a breast cancer incidence which is about ¼ of that in the United States. We show that the estimated disease risk given a particular family history is lower in Japan than in the United States or in Germany, although the relative difference is not very large.
Conclusion: The proposed method can be used to generate tables which may be useful for genetic counselling in arbitrary populations.

66

Association Between Family History of Cancer and Breast Tumor Estrogen and Progesterone Receptor Status. AM Tutera, CR Drinkard, JD Potter, GL Wiesner, AR Folsom, TA Sellers

We recently reported a modest inverse association of family history with estrogen receptor negative (ER-) /progesterone receptor positive (PR+) breast cancer in the Iowa Women's Health Study, a cohort study of 41,837 postmenopausal women. However, that analysis did not consider 2nd degree relatives or cancer sites that may be etiologically related. Family histories of breast, ovarian, uterine, cervical, or non-specified reproductive organs in 1st and 2nd degree relatives were collected in 1986. Data on prostate cancer in fathers and brothers were collected in 1992. Cohort members were followed for cancer incidence through the state-wide tumor registry. After 7 years of follow-up, 939 breast cancers were identified; 65.0% had information on ER (+/-) and PR (+/-) status. A family history of breast cancer in 1st degree relatives was associated with increased risk (RR = 1.37; 95% CI: 1.14 - 1.63) for all receptor defined subtypes of breast cancer except ER+/PR- tumors (RR = 0.66; 95% CI: 0.32 - 1.36). Similar results were seen for ER+/PR- tumors when breast cancer among 2nd degree relatives was considered (RR = 0.77; 95% CI: 0.40 - 1.49). A family history of reproductive/breast cancer in 1st degree relatives was associated with a non-statistically significant increased risk for ER-/PR+ breast cancer (RR = 2.24; 95% CI: 0.83 - 6.66) but no other subtypes. A family history of prostate cancer in first-degree relatives was associated with an 1.22-fold increased risk of breast cancer (95% CI: 0.98 - 1.50), largely a reflection of the association with ER-PR- tumors (RR = 1.49; 95% CI: 0.76-2.95). The small numbers of cases in some categories and the corresponding wide confidence intervals preclude definitive conclusions, but these data are at least suggestive that joint stratification of breast tumors on ER and PR status may be useful in partioning breast cancer families into more homogeneous subsets.