

Computer Applications in Genetic Epidemiology

Spring 2011

Instructor: Karen Edwards, PhD (Professor)
Email: keddy@uw.edu
Office: F363E, Health Sciences Center
Office Hours: By Appointment
Phone: 206-616-1258

Meeting Times and Locations

Tuesdays and Thursdays

11:00 am - 12:50 pm

Health Sciences Center E216*

*Class will occasionally meet in a Microcomputer Lab in the Health Sciences Library

Course Description

This course will provide an overview of genetic epidemiologic analyses and demonstrate the use of specialized software used for these analyses. In particular, analyses will include heritability estimation in segregation and linkage analyses in family-based studies, haplotype estimation and analysis in association studies with unrelated individuals, and power of genetic association studies. This course will incorporate examples from the literature, in-class demonstrating as well as computer lab exercises of analyses. Students will have the opportunity to perform analyses using data sets supplied by the instructor and will present their results in class.

Grading Policy

Four Credit Option:

Project - 50%

Homework - 30%

Class Participation - 20%

Two Credit Option:

Homework - 80%

Class Participation - 20%

Disability Accommodations

If you would like to request academic accommodations due to a disability, please contact:

| | |
|--|---|
| Disability Resources for Students 448 Schmitz Hall Campus Box 355839 Seattle, WA 98195-5839 | Phone: 206-543-8924 (V); 206-543-8925 (V/TTY) Fax: 206-616-8379 Email: uwdss@uw.edu |
|--|---|

If you have a letter from the Office of Disability Resources for Students indicating that you have a disability which requires academic accommodations, please present the letter to the instructor so accommodations may be made.

PHG 518 Spring 2011 Schedule

| Week/Date | Lecture Topic | Format | Work Assigned | Work Due |
|------------------------|--|--|--|----------------------------------|
| Week 1 | | | | |
| Tuesday – March 29 | Course Overview Introduction | Discussion | None | None |
| Thursday – March 31 | Commingling Analysis Twin Studies | Discussion & Demo | Homework #1: Commingling Analysis | None |
| Week 2 | | | | |
| Tuesday – April 5 | Collecting Family Data Familial Correlations | Discussion | Homework #2: Pedcheck | <i>HW #1</i> |
| Thursday – April 7 | Family Data cont. Pedigree Drawing Pedigree Checking | Discussion & Demo | None | None |
| Week 3 | | | | |
| Tuesday – April 12 | Complex Segregation Analysis | Discussion | None | None |
| Thursday – April 14 | Complex Segregation Analysis | Discussion | Homework #3: Segregation Analysis | None |
| Week 4 | | | | |
| Tuesday – April 19 | Quantitative Genetic Analysis | Discussion | None | <i>HW #2</i> |
| Thursday – April 21 | Linkage | Discussion | Homework #4: Linkage Analysis | <i>HW #3</i> |
| Week 5 | | | | |
| Tuesday – April 26 | Solar Linkage Analysis | Discussion Meet in HSL | None | None |
| Thursday – April 28 | Solar Linkage Analysis | Demo Meet in HSL | Data Analysis Project: Part #1 | <i>HW #4</i> |
| Week 6 | | | | |
| Tuesday – May 3 | Hardy Weinberg Equilibrium Genetic Association Population Stratification | Discussion | None | None |
| Thursday – May 5 | Transmission Disequilibrium Test | Discussion | None | None |
| Week 7 | | | | |
| Tuesday – May 10 | Linkage Disequilibrium Haplotypes Genome Wide Association Studies | Discussion | None | None |
| Thursday – May 12 | PLINK GWAS Analysis | Demo Meet in HSL | None | None |
| Week 8 | | | | |
| Tuesday – May 17 | Work on DAP | Lab Meet in HSL | Data Analysis Project: Part #3 | None |
| Thursday – May 19 | Genome Variation Server | Guest Lecturer: Dr. Roberston Meet in HSL | None | None |
| Week 9 | | | | |
| Tuesday – May 24 | Work on DAP | Lab Meet in HSL | Data Analysis Project: Part #2 | None |
| Thursday – May 26 | Post-GWAS Analysis | Guest Lecturer: Dr. Mata Meet in HSL | None | None |
| Week 10 | | | | |
| Tuesday – May 31 | Power Wrap-up | Discussion | None | None |
| Thursday – June 2 | Student Presentations | Presentations | None | Data Analysis Project |

Reading Assignments

Week 1: March 29, March 31

Tuesday March 29

Course Overview / Intro to Genetic Epi

- 1) [Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. Lancet. Sep 10-16 2005;366\(9489\):941-951.](#)
- 2) [Manolio TA, Collins FS, Cox NJ, et. al., Finding the missing heritability of complex diseases. Nature 2009; 461\(8\): 747-753.](#)

Thursday March 31

A) Commingling Analysis / NOCOM

- 1) [Kardia SLR, Sing CF, Turner ST. The response of renal plasma flow to angiotensin II infusion in a population-based sample and its association with parental history of essential hypertension. J of Hypertension 1997;15:483-93.](#)
- 2) [Mayer EJ, Newman B, Austin M, et al., Genetic and environmental influences on insulin levels and the IRS: an analysis of women twins. Am J Epidemiol 1996; 143:323-32.](#)

B) Twin Studies

- 1) [Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. Nat Rev Genet. 2002 Nov;3\(11\):872-82.](#)
- 2) [Goldberg J, True WR, Eisen SA, and Henderson WG. A twin study of the effects of the Vietnam War on posttraumatic stress disorder. JAMA. March 2, 1990.](#)
- 3) [True WR, Rice J, Eisen SA, et al. A twin study of genetic and environmental contributions to liability for posttraumatic stress symptoms. Arch Gen Psychiatry. 50:257-264, 1993.](#)

C) Supplemental Readings

- 1) [Zhao LP et al. Population-based family study designs: an interdisciplinary research framework for genetic epidemiology. Genetic Epidemiol. 1997;14:365-88.](#)
- 2) [Kwon JW, Boehnke M, Burns TL, Moll PP. Commingling and Segregation analyses: comparison of results from a simulation study of a quantitative trait. Genetic Epi 1990;7:57-68.](#)

Week 2: April 5, April 7

Tuesday April 5 - **HW1 Due**

A) Collecting Family Data and ELSI Issues

- 1) [Botkin, JR, McMahon WM, Smith KR, Nash JE. Privacy and confidentiality in the publication of pedigrees. JAMA 1998;279:1808-12](#)
- 2) [Holtzman NA, Andrews LB. Ethical and legal issues in genetic epidemiology. Epidemiol Rev. 1997;19:163-174](#)
- 3) [Botkin JR. Protecting the privacy of family members in survey and pedigree research. JAMA. 2001;285:207-11](#)
- 4) [Byers PH, Ashkenas J. Pedigrees-Publish? Or Perish the Thought? Am J Hum Genet. 1998;63:678-81](#)

B) Family Data Cleaning

- 1) [Bennett RL, French KS, Resta RG, Doyle DL. Standardized Human Pedigree Nomenclature: Update and Assessment of the Recommendations of the National Society of Genetic Counselors. J Genet Counsel 2008;17:424-433](#)

C) Supplemental Readings

- 1) [Bennett RL, Steinhaus KA, Uhrich SB, et al., Recommendations for standardized human pedigree nomenclature. Am J Hum Genet 1995;56:745-752](#)
- 2) [Cannings C, Thompson EA. Ascertainment in the sequential sampling of pedigrees. Clinical Genetics 1977;12:208-212](#)

Thursday April 7

Family Data

No Readings Assigned

Week 3: April 12, April 14

Tuesday April 12

Familial Aggregation / Complex Segregation Analysis

- 1) [Pankow JS, Folsom AR, Cushman M, Borecki IB, Hopkins PN, Eckfeldt JH, Trace RP. Familial and genetic determinants of systemic markers of inflammation: the NHLBI family heart study. Atherosclerosis 2001;154:681-9.](#)
- 2) [Pairitz G, Davignon J, Mailloux H, Sing CF. Sources of interindividual variation in the quantitative levels of apolipoprotein B in pedigrees ascertained through a lipid clinic Am J Hum Genet. 1988 Sep;43\(3\):311-21.](#)
- 3) [Jarvik G. Complex Segregation Analysis: Uses and Limitations. Am J Hum Genetic 1998;63:942-6.](#)
- 4) [Schumacher MC, Hasstedt SJ, Hunt SC, et al. Major gene effect for insulin levels in familial NIDDM pedigrees. Diabetes 1992;41:416-423](#)

Thursday April 14

Complex Segregation Analysis / Quantitative Genetic Analysis

[Borecki IB, Province MA, Rao DC. Inferring a major gene for quantitative traits by using segregation analysis with tests on transmission probabilities: How often do we miss? Am J Hum Genetic 1995;56:319-326.](#)

Week 4: April 19, April 21

Tuesday April 19 - **HW2 Due**

- 1) [Edwards KL, Mahaney MC, Motulsky AG, Austin MA. Pleiotropic genetic effects on LDL size, plasma triglyceride, and HDL cholesterol in families. Arterioscler Thromb Vasc Biol. 1999 Oct;19\(10\):2456-64.](#)
- 2) [Edwards KL, Hutter CM, Wan JY, Kim H, Monks SA. Genome-wide linkage scan for the metabolic syndrome: the GENNID study. Obesity. 2008 Jul;16\(7\):1596-601.](#)
- 3) [Edwards KL, Wan JY, Hutter CM, Fong PY, Santorico SA. Multivariate Linkage Scan for Metabolic Syndrome Traits in Families with Type 2 Diabetes. Obesity. \[Epub ahead of print\]](#)

Thursday April 21 - **HW3 Due**

Linkage

- 1) [Elston RC. Methods of linkage analysis--and the assumptions underlying them. Am J Hum Genet. 1998 Oct;63\(4\):931-4.](#)
- 2) [Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, Balow J Jr, Ally DS, Kohtamaki K, Chines P, Birznieks G, Kaleta HS, Musick A, Te C, Tannenbaum J, Eldridge W, Shapiro S, Martin C, Witt A, So A, Chang J, Shurtleff B,](#)

[Porter R, Boehnke M, et al. Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. Proc Natl Acad Sci U S A. 1999 Mar 2;96\(5\):2198-2203.](#)

3) [Hokanson JE, Brunzell JD, et al. Linkage of LDL size to the lipoprotein lipase gene in heterozygous lipoprotein lipase deficiency. Am J Hum Genet 1999;64.](#)

Week 5: April 26, April 28

Tuesday April 26**Meet in Computer Lab in HSL**

Solar Linkage Analysis

1) Look through tutorial of SOLAR: <http://solar.sfbgenetics.org/doc/03.chapter.html> (Note: This tutorial will be demonstrated in-class in the computer lab.)

2) [Bellis C, et al. Linkage Mapping of CVD Risk Traits in the Isolated Norfolk Island Population. Hum Genet \(2008\) 124:543-552.](#)

3) [Wu J, et al. A QTL on 12q Influencing an Inflammation Marker and Obesity in White Women: The NHLBI Family Heart Study. Obesity \(Silver Spring\). 2009 Mar;17\(3\):525-31. Epub 2008 Dec 25](#)

Thursday April 28 - **HW4 Due**

Meet in Computer Lab in HSL

Solar Linkage Analysis

None Assigned

Week 6: May 3, May 5

Tuesday May 3

A) Hardy Weinberg Equilibrium and Association Studies

1) [Khoury MJ, Beaty TH. Applications of the case-control method in genetic epidemiology. Epidemiol Rev. 1994;16:134-150.](#)

2) [Yang Q, Khoury MJ. Evolving methods in genetic epidemiology: III. Gene-environment interaction in epidemiological research. Epidemiol Rev 1997;19:22-43](#)

3) [Schaid DJ & Jacobsen SJ. Biased tests of association: Comparison of allele frequencies when departing from Hardy-Weinberg proportions. AJE. 1999;149:706-11](#)

4) [Cordell HJ, Clayton DG. Genetic Epidemiology 3: Genetic association studies. The Lancet; Sep 24-30, 2005; 366: 1121-1131.](#)

5) [Love-Gregory LD, Wasson J, Ma J, Jin CH, Glaser B, Suarez BK, Permutt MA. A Common Polymorphism in the Upstream Promoter Region of the HNF4 Gene on Chromosome 20q Is Associated With Type 2 Diabetes and Appears to Contribute to the Evidence for Linkage in an Ashkenazi Jewish Population. Diabetes. 2004;53:1134-40.](#)

6) [Silander K, Mohlke KL, Scott LJ, Peck EC, et al. Genetic Variation Near the Hepatocyte Nuclear Factor-4HNF \$\alpha\$ Gene Predicts Susceptibility to Type 2 Diabetes. Diabetes. 2004;53\(4\):1141-1150.](#)

B) Population Stratification

7) [Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet. 2003 Feb 15;361\(9357\):598-604.](#)

8) [Knowler et al. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture Am J Hum Genet. 1988 Oct;43\(4\):520-6.](#)

9) [Pharoah PD, Tyrer J, Dunning AM, Easton DF, Ponder BA; SEARCH Investigators. Association between Common Variation in 120 Candidate Genes and Breast Cancer Risk. PLoS Genet. 2007 Mar 16;3\(3\):e42](#)

Thursday May 5

Family-based tests of association

- 1) [Schaid DJ. Transmission disequilibrium, family controls, and great expectations. Am J Hum Genet. 1998; 63:935-41.](#)
- 2) [Beatty TH, et al. Testing candidate genes for non-syndromic oral clefts using a case-parent trio design. Gen Epidemiol 2002;22\(1\):1-11.](#)
- 3) [Laird C, Lange NM. Family-based designs in the age of large-scale gene-association studies. Nature Reviews Genetics. 2006; 7: 385-394.](#)

Week 7: May 10, May 12

Tuesday May 10

A) Association Studies

- 1) [Zondervan KT and Cardon LR. Designing candidate gene and genome-wide case-control association studies. Nature Protocols 2007; 2\(10\): 2492-501.](#)
- 2) [Bhatia G, Bansal V, Harismendy O, Schork NJ, Topol EJ, Frazer K, Bafna V. A covering method for detecting genetic associations between rare variants and common phenotypes. PLoS Comput Biol 2010; 6\(10\): e1000954.](#)
- 3) [Thomas D. Methods for Investigating Gene-Environment Interactions in Candidate Pathway and Genome-Wide Association Studies. Annu Rev Public Health 2010; 31\(4\): 21-36.](#)
- 4) [Thomas D. Gene-Environment-Wide Association Studies: Emerging Approaches. Nat Rev Gen 2010; 11\(4\): 259-72.](#)

B) HapMap

- 2) [The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007 \(449\): 851-861.](#)

C) Linkage Disequilibrium

- 3) [Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. Nat Rev Genet 2004 \(5\): 89-100.](#)

D) Haplotypes

- 4) [Schaid DJ, McDonnell SK, Wang L, Cunningham JM, Thibodeau SN. Caution on pedigree haplotype inference with software that assumes linkage equilibrium. Am J Hum Genet. 2002 Oct;71\(4\):992-5.3.](#)
- 5) [Lin DY and Huang BE. \(2007\) The use of inferred haplotypes in downstream analyses. Am J Hum Genet, 80:577-579.](#)

E) Supplemental Readings

- 1) [Lewontin RC. On measures of gametic disequilibrium. Genetics. 1988 Nov;120\(3\):52.](#)
- 2) [McVean G, Spencer CC, Chaix R. Perspectives on human genetic variation from the HapMap Project. PLoS Genet. 2005 Oct;1\(4\):e54](#)
- 3) [Carlson CS, Heagerty PJ, Nord AS, et al. TagSNP evaluation for the association of 42 inflammation loci and vascular disease: evidence of IL6, FGB, ALOX5, NFKBIA, and IL4R loci effects Human Genetics 2007 Mar; 121 \(1\):65-75.](#)

4) [Hindorff LA, Psaty BM, Carlson CS, Heckbert SR, Lumley T, et al. \(2006\) Common genetic variation in the prothrombin gene, hormone therapy, and incident nonfatal myocardial infarction in postmenopausal women. American Journal of Epidemiology 163\(7\):600-607.](#)

5) [Carty, CL, Cushman M, et al \(2008\) Associations between common fibrinogen gene polymorphisms and cardiovascular disease in older adults. The Cardiovascular Health Study. Thromb Haemost. 99\(2\):388-95.](#)

Thursday May 12Meet in Computer Lab in HSL****

PLINK Tutorial
None Assigned

Week 8: May 17, May 19

Tuesday May 17Meet in Computer Lab in HSL****

GWAS Analysis
None Assigned

Thursday May 19Meet in Computer Lab in HSL****

Guest Lecturer: Dr. Peggy Robertson on the Genome Variation Server
None Assigned

Week 9: May 24, May 26

Tuesday May 24Meet in Computer Lab in HSL****

Tag SNP Selection
None Assigned

Thursday May 26Meet in Computer Lab in HSL****

Guest Lecturer: Dr. Ignacio Mata on Post-GWAS Analysis
None Assigned

Week 10: May 31, June 2

Tuesday May 31

A) Power

1) [Gauderman WJ. Sample size requirements for matched case-control of gene-environment interaction. Stat Med 21:35-50, 2002.](#)

2) [Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol 155:478-84, 2002.](#)

3) [Gauderman WJ. Candidate gene association studies for a quantitative using parent-offspring trios. Genet Epidemiol 25:327-338, 2003.](#)

4) [Dupont WD, Plummer WD, Jr: Power and Sample Size Calculations: A Review and Computer Program. Controlled Clinical Trials 11:116-128, 1990](#)

5) [Dupont WD, Plummer WD, Jr: Power and Sample Size Calculations for Studies Involving Linear Regression. Controlled Clinical Trials 19:589-601, 1998](#)

6) [Lange C, Demeo D, Silverman EK, Weiss ST, Laird DM. \(2004\) PBAT: tools for family-based association studies. Am J Hum Genet. 74\(2\):367-9.](#)

B) Power Programs

7) Quanto - Gauderman WJ, Morrison JM. QUANTO 1.2.3: A computer program power and sample size calculations for genetic-epidemiology studies, <http://hydra.usc.edu/gxe>, May 2007.

8) [Genetic Power Calculator - Purcell S, Cherny SS, Sham PC. \(2003\) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics, 19\(1\):149-150.](#)

Thursday June 2

Student Presentations

None Assigned

Homework

Please turn in a hard copy of your homework in class. If you are out of town and unable to turn in your work during class, you can submit an electronic copy of your homework to the drop box.

Homework #1: Commingling Analysis

Homework #2: Pedcheck

Homework #3: Segregation Analysis

Homework #4: Linkage Analysis

Homework #5: Hardy Weinberg Equilibrium and Family Based Association

Homework #6: Haplotype Measures and Analysis & Power

Data Analysis Project

For students taking this course for 4 credits, a data analysis project is required. The data will be supplied by the instructor, and students will present their results in class during finals week.

Overview

The data analysis project (DAP) is comprised of the following three parts:

Part 1: SOLAR VC linkage analysis

Each student will be given a different dataset to analyze using SOLAR linkage software. They will also be required to answer interpretation questions based on their linkage results.

Part 2: Tag SNP selection

Each student will select tagSNPs for a gene of their choice. For this portion of the assignment, students should write up a brief description of the process which could be used in a methods section for a manuscript or grant proposal. They will also answer interpretation questions.

Part 3: Genetic association analysis

Each student will be provided with a dataset to perform a genome-wide association analysis. The results of the genome-wide analysis and a haplotype analysis should be written up and interpreted.