

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/35195> holds various files of this Leiden University dissertation

**Author:** Balliu, Brunilda

**Title:** Statistical methods for genetic association studies with response - selective sampling designs

**Issue Date:** 2015-09-10

# Bibliography

- Ahlbach C, Usatine J, and Pippenger N 2012. A combinatorial interpretation of the joint cumulant. arXiv:12110652 [math] arXiv: 1211.0652.
- Ainsworth HF, Unwin J, Jamison DL, and Cordell HJ 2011. Investigation of maternal effects, maternal-fetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring. *Genet Epidemiol* 35, no. 1:19–45.
- Akey J, Jin L, and Xiong M 2001. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J Hum Genet* 9, no. 4:291–300.
- Azzouz D, Martin M, Roudier J, and Lambert NC 2011. Could microchimerism be a source of disease-associated HLA alleles in patients with scleroderma? *Ann Rheum Dis* 70:A34.
- Bennett JH 1952. On the theory of random mating. *Annals of Eugenics* 17, no. 1:311–317.
- Boehringer S, Guenther M, Sinigerova S, Wurtz RP, Horsthemke B, and Wieczorek D 2011. Automated syndrome detection in a set of clinical facial photographs. *American Journal of Medical Genetics Part A* 155, no. 9:2161–2169.
- Boehringer S and Pfeiffer RM 2009. A model for fine mapping in family based association studies. *Hum Hered* 67, no. 4:226–236.
- Boehringer S, Vollmar T, Tasse C, Wurtz RP, Gillessen-Kaesbach G, Horsthemke B, and Wieczorek D 2006. Syndrome identification based on 2d analysis software. *European Journal of Human Genetics: EJHG* 14, no. 10:1082–9.
- Breslow NE and Day NE 1980. *Statistical methods in cancer research , 1: The analysis of case-control studies*. Lyon: International Agency for Research on Cancer .
- Breslow NE, Robins JM, and Wellner JA 2000. On the semi-parametric efficiency of logistic regression under case-control sampling. *Bernoulli* 6:447–455.
- Brillinger DR 1991. Some history on the study of higher order moments and spectra. *Statistica Sinica* 1 1:465–476.
- Brown MD, Glazner CG, Zheng C, and Thompson EA 2012. Inferring coancestry in population samples in the presence of linkage disequilibrium. *Genetics* 190, no. 4:1447–1460.
- Browning BL and Browning SR 2011. A fast, powerful method for detecting identity by descent. *Am J Hum Genet* 88, no. 2:173–182.
- Browning SR and Thompson EA 2012. Detecting rare variant associations by identity-by-descent mapping in case-control studies. *Genetics* 190, no. 4:1521–1531.
- Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, . . . , and Compston A 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, no. 7145:661–678.

- Calza S and Pawitan Y 2010. Normalization of gene-expression microarray data. *Methods Mol Biol* pp. 37–52.
- Cantor RM, Lange K, and Sinsheimer JS 2010. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *Am J Hum Genet* 86, no. 1:6–22.
- Cardon LR and Palmer LJ 2003. Population stratification and spurious allelic association. *Lancet* 361, no. 9357:598–604.
- Chatterjee N and Carroll RJ 2005. Semiparametric maximum likelihood estimation exploiting gene-environment independence in case-control studies. *Biometrika* 92:399–418.
- Chen J, Lin D, and Hochner H 2012. Semiparametric maximum likelihood methods for analyzing genetic and environmental effects with case-control mother-child pair data. *Biometrics* pp. 10.1111/j.1541-0420.2011.01728.x.
- Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U, and Hsu L 2010. Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *Am J Hum Genet* 86, no. 6:860–871.
- Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, Zhang C, Lamb J, Edwards S, Sieberts SK, Leonardson A, Castellini LW, Wang S, Champy MF, Zhang B, Emilsson V, Doss S, Ghazalpour A, Horvath S, Drake TA, Lusk AJ, and Schadt EE 2008. Variations in DNA elucidate molecular networks that cause disease. *Nature* 452:429–435.
- Clark AG, Weiss KM, Nickerson DA, Taylor SL, Buchanan A, Stengård J, Salomaa V, Vartiainen E, Perola M, Boerwinkle E, and Sing CF 1998. Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase. *Am J Hum Genet* 63, no. 2:595–612.
- Clayton D and Jones H 1999. Transmission/disequilibrium tests for extended marker haplotypes. *The American Journal of Human Genetics* 65, no. 4:1161–1169.
- Degli-Esposti MA, Leelayuwat C, and Dawkins RL 1992. Ancestral haplotypes carry haplotypic and haplospecific polymorphisms of BAT1: possible relevance to autoimmune disease. *Eur J Immunogenet* 19, no. 3:121–127.
- Dempster AP, Laird NM, and Rubin DB 1977. Maximum likelihood from incomplete data via the EM algorithm. *J R Statist Soc B* 39, no. 1:1–38.
- Diggle Pj, Liang KY, and Zeger SL 1994. *Analysis of longitudinal data*. Oxford Science Publications.
- Dimas AS, Nica AC, Montgomery SB, Stranger BE, Raj T, Buil A, Giger T, Lappalainen T, Gutierrez-Arcelus M, McCarthy MI, and Dermitzakis ET 2012. Sex-biased genetic effects on gene regulation in humans. *Genome Res* 22:2368–2375.
- Drysdale CM, McGraw DW, Stack CB, Stephens JC, Judson RS, Nandabalan K, Arnold K, Ruano G, and Liggett SB 2000. Complex promoter and coding region beta 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. *Proc Natl Acad Sci USA* 97, no. 19:10483–10488.
- Edgar R, Domrachev M, and Lash AE 2002. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucl Acids Res* 30:207–210.

- Epstein MP and Satten GA 2003. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 73, no. 6:1316–1329.
- Evangelou E, Trikalinos TA, Salanti G, and Ioannidis JPA 2006. Family-based versus unrelated case-control designs for genetic associations. *PLoS Genet* 2, no. 8:e123.
- Excoffier L and Slatkin M 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12, no. 5:921–927.
- Feitsma AL, Worthington J, van der Helm-van Mil AHM, Plant D, Thomson W, Ursum J, van Schaardenburg D, van der Horst-Bruinsma IE, van Rood J, Huizinga TWJ, Toes REM, and de Vries RRP 2007. Protective effect of noninherited maternal HLA-DR antigens on rheumatoid arthritis development. *Proc Natl Acad Sci USA* 104:19966–70.
- Fisher R 1930. *The Genetical Theory of Natural Selection*. Oxford University Press.
- Fitzmaurice BY, Garrett M, and Laird N 1993. A likelihood-based method for analysing longitudinal binary responses. *Biometrika* 80:141–51.
- Friedman J, Hastie T, and Tibshirani R 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software* 33, no. 1:1–22.
- Glass D, Viñuela A, Davies MN, Ramasamy A, Parts L, Knowles D, Brown AA, Hedman AK, Small KS, Buil A, Grundberg E, Nica AC, Meglio PD, Nestle FO, Ryten M, the UK Brain Expression consortium, the MuTHER consortium, Durbin R, McCarthy MI, Deloukas P, Dermitzakis ET, Weale ME, Bataille V, and Spector TD 2013. Gene expression changes with age in skin, adipose tissue, blood and brain. *Genome Biology* 14:R75.
- Gower JC 1975. Generalized procrustes analysis. *Psychometrika* 40, no. 1:33–51.
- Günther M 2012. *Statistical Gabor graph based techniques for the detection, recognition, classification, and visualization of human faces*. Ph.D. thesis, Shaker, Aachen.
- Gutierrez-Arcelus M, Lappalainen T, Montgomery SB, Buil A, Ongen H, Yurovsky A, Bryois J, Giger T, Romano L, Planchon A, Falconnet E, Bielser D, Gagnebin M, Padioleau I, Borel C, Letourneau A, Makrythanasis P, Guipponi M, Gehrig C, Antonarakis SE, and Dermitzakis ET 2013. Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *eLife Sciences* 2:e00523.
- Hammond P, Hannes F, Suttie M, Devriendt K, Vermeesch JR, Faravelli F, Forzano F, Parekh S, Williams S, McMullan D, South ST, Carey JC, and Quarrell O 2012. Fine-grained facial phenotype-genotype analysis in wolf-hirschhorn syndrome. *European Journal of Human Genetics* 20, no. 1:33–40.
- Hammond P, Hutton TJ, Allanson JE, Buxton B, Campbell LE, Clayton-Smith J, Donnai D, Karmiloff-Smith A, Metcalfe K, Murphy KC, Patton M, Pober B, Prescott K, Scambler P, Shaw A, Smith ACM, Stevens AF, Temple IK, Hennekam R, and Tassabehji M 2005. Discriminating power of localized three-dimensional facial morphology. *The American Journal of Human Genetics* 77, no. 6:999–1010.
- Hastie T, Tibshirani R, and Friedman J 2001. *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York, USA.
- Hattersley AT and McCarthy MI 2005. What makes a good genetic association study? *Lancet* 366, no. 9493:1315–1323.

- Hay EM, Olliver WFR, and Silman AJ 1993. The arthritis and rheumatism council's national family material repository. *Br J Rheumatol* 32:443–444.
- Heagerty PJ and Kurland BF 2001. Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika* 88:973–985.
- Hedrick PW 1987. Gametic disequilibrium measures: Proceed with caution. *Genetics* 117, no. 2:331–341.
- Hennessy RJ, Baldwin PA, Browne DJ, Kinsella A, and Waddington JL 2007. Three-dimensional laser surface imaging and geometric morphometrics resolve frontonasal dysmorphology in schizophrenia. *Biological Psychiatry* 61, no. 10:1187–1194.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, and Manolio TA 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106:9362–9367.
- Hollox EJ, Poulter M, Zvarik M, Ferak V, Krause A, Jenkins T, Saha N, Kozlov AI, and Swallow DM 2001. Lactase haplotype diversity in the old world. *Am J Hum Genet* 68, no. 1:160–172.
- Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, van den Berg LH, and Ophoff RA 2012. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol* 13:R97.
- Houwing-Duistermaat JJ, Helmer Q, Balliu B, Akker Evd, Tsonaka R, and Uh HW 2014. Gene analysis for longitudinal family data using random-effects models. *BMC Proceedings* 8, no. Suppl 1:S88.
- Houwing-Duistermaat JJ, van Houwelingen HC, and de Winter J P 2000. Estimation of individual genetic effects from binary observations on relatives applied to a family history of respiratory illnesses and chronic lung disease of newborns. *Biometrics* 56, no. 3:808–14.
- Howden L and Meyer J 2011. Age and sex composition: 2010. US Department of Commerce, Economics and Statistics Administration, US Census Bureau .
- Hsieh H, Palmer CGS, Harney S, Chen H, Bauman L, Brown MA, and Sinsheimer JS 2007. Using the maternal-fetal genotype incompatibility test to assess non-inherited maternal HLA-DRB1 antigen coding alleles as rheumatoid arthritis risk factors. *BMC Proc* 1(Suppl 1):S124.
- Hsieh H, Palmer CGS, Harney S, Newton JL, Wordsworth P, Brown MA, and Sinsheimer JS 2006. The v-MFG test: Investigating maternal, offspring and maternal-fetal genetic incompatibility effects on disease and viability. *Genet Epidemiol* 30:333–347.
- Huang YT, VanderWeele TJ, and Lin X 2014. Joint analysis of SNP and gene expression data in genetic association studies of complex diseases. *The Annals of Applied Statistics* 8:352–376.
- Huijts PEA, Hollestelle A, Balliu B, Houwing-Duistermaat JJ, Meijers CM, Blom JC, Ozturk B, Krol-Warmerdam EMM, Wijnen J, Berns EMJJ, Martens JWM, Seynaeve C, Kiemeny LA, van der Heijden HF, Tollenaar RAEM, Devilee P, and van Asperen CJ 2014. CHEK2\*1100delC homozygosity in the netherlands—prevalence and risk of breast and lung cancer. *Eur J Hum Genet* 22:46–51.

- Huynh JL, Garg P, Thin TH, Yoo S, Dutta R, Trapp BD, Haroutunian V, Zhu J, Donovan MJ, Sharp AJ, and Casaccia P 2014. Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat Neurosci* 17:121–130.
- Ioannidis JPA 2003. Genetic associations: false or true? *Trends Mol Med* 9, no. 4:135–138.
- Jolliffe I 2005. Principal component analysis. Wiley Online Library.
- Joosten PH, Toepoel M, Mariman EC, and Van Zoelen EJ 2001. Promoter haplotype combinations of the platelet-derived growth factor alpha-receptor gene predispose to human neural tube defects. *Nat Genet* 27, no. 2:215–217.
- Kenny DA and Judd CM 2013. Power anomalies in testing mediation. *Psychological Science* 25:334–339.
- Koch L 2014. Epigenetics: An epigenetic twist on the missing heritability of complex traits. *Nat Rev Genet* 15, no. 4:218–218.
- Kraft P, Hsieh HJ, Cordell HJ, and Sinsheimer J 2005. A conditional-on-exchangeable-parental-genotypes likelihood that remains unbiased at the causal locus under multiple-affected-sibling ascertainment. *Genet Epidemiol* 29, no. 1:87–90.
- Kraft P and Thomas DC 2000. Bias and efficiency in family-based gene-characterization studies: conditional, prospective, retrospective, and joint likelihoods. *Am J Hum Genet* 66:1119–31.
- Lewontin RC 1964. The interaction of selection and linkage. i. general considerations; heterotic models. *Genetics* 49, no. 1:49–67.
- Li H 2013. Systems biology approaches to epidemiological studies of complex diseases. *WIREs Syst Biol Med* 5:677–686.
- Liang KY and Zeger SL 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73:13–22.
- Lin DY and Zeng D 2006. Likelihood-based inference on haplotype effects in genetic association studies. *Journal of the American Statistical Association* 101, no. 473:89–104.
- Lin DY and Zeng D 2009. Proper analysis of secondary phenotype data in case-control association studies. *Genet Epidemiol* 33:256–265.
- Liu J, Morgan M, Hutchison K, and Calhoun VD 2010. A study of the influence of sex on genome wide methylation. *PLoS ONE* 5:e10028.
- Liu N, Zhang K, and Zhao H 2008. Haplotype-association analysis. In: *Advances in Genetics*, Academic Press, volume 60 of *Genetic Dissection of Complex Traits*, pp. 335–405.
- Lv J, Liu H, Su J, Wu X, Liu H, Li B, Xiao X, Wang F, Wu Q, and Zhang Y 2012. DiseaseMeth: a human disease methylation database. *Nucleic Acids Res* 40 (Database issue):D1030–1035.
- Mefford J and Witte JS 2012. The covariate's dilemma. *PLoS genetics* 8:e1003096.
- Meinshausen N 2008. Hierarchical testing of variable importance. *Biometrika* 95, no. 2:265–278.

- Molinaro AM, Simon R, and Pfeiffer RM 2005. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 21, no. 15:3301–3307.
- Morris RW and Kaplan NL 2002. On the advantage of haplotype analysis in the presence of multiple disease susceptibility alleles. *Genet Epidemiol* 23, no. 3:221–233.
- Mukherjee B and Chatterjee N 2008. Exploiting gene-environment independence for analysis of case-control studies: an empirical bayes-type shrinkage estimator to trade-off between bias and efficiency. *Biometrics* 64:685–694.
- Nelson JL, Furst DE, Maloney S, Gooley T, Evans PC, Smith A, Bean MA, Ober C, and Bianchi D 1998. Microchimerism and HLA-compatible relationships of pregnancy in scleroderma. *Lancet* 351:559–62.
- Nemes S, Jonasson JM, Genell A, and Steineck G 2009. Bias in odds ratios by logistic regression modelling and sample size. *BMC Med Res Methodol* 9:56.
- Neuhaus JM 1998. Estimation efficiency with omitted covariates in generalized linear models. *Journal of the American Statistical Association* 93:1124–1129.
- Neuhaus JM and Jewell NP 1993. A geometric approach to assess bias due to omitted covariates in generalized linear models. *Biometrika* 80:807–815.
- Pfeiffer RM, Hildesheim A, Gail MH, Pee D, Chen CJ, Goldstein AM, and Diehl SR 2003. Robustness of inference on measured covariates to misspecification of genetic random effects in family studies. *Genet Epidemiol* 24:14–23.
- Pfeiffer RM, Pee D, and Landi MT 2008. On combining family and case-control studies. *Genet Epidemiol* 32:638–46.
- Pirinen M, Donnelly P, and Spencer CCA 2012. Including known covariates can reduce power to detect genetic effects in case-control studies. *Nature Genetics* 44:848–851.
- Prentice RL and Pyke R 1979. Logistic disease incidence models and case-control studies. *Biometrika* 66:403–411.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, and Reich D 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38, no. 8:904–909.
- Purdom E and Holmes SP 2005. Error distribution for gene expression data. *Statistical applications in genetics and molecular biology* 4:1544–6115.
- R Core Team 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rabinowitz D 1997. A note on efficient estimation from case-control data. *Biometrika* 84:486–488.
- Richardson B 2003. Impact of aging on DNA methylation. *Ageing Research Reviews* 2:245–261.
- Risch N and Merikangas K 1996. The future of genetic studies of complex human diseases. *Science* 273, no. 5281:1516–1517.

- Robinson LD and Jewell NP 1991. Some surprising results about covariate adjustment in logistic regression models. *International Statistical Review / Revue Internationale de Statistique* 59:227–240.
- Schadt EE, Monks SA, Drake TA, Lusk AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, and Friend SH 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422:297–302.
- Schaid DJ 2004. Evaluating associations of haplotypes with traits. *Genet Epidemiol* 27, no. 4:348–364.
- Schaid DJ, Rowland CM, Tines DE, Jacobson RM, and Poland GA 2002. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 70, no. 2:425–434.
- Scott AJ and Wild CJ 2001. Maximum likelihood for generalised case-control studies. *Journal of Statistical Planning and Inference* 96:3–27.
- Silman AJ and Hochberg MC, editors 2001. *Epidemiology of the Rheumatic Diseases*. Oxford ; New York: Oxford University Press, 2 edition edition.
- Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, and Ollier WER 1993. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol* 32:903–7.
- Sinnwell J and Schaid D 2013. *haplo.stats: Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous*. R package version 1.6.8.
- Sinsheimer JS, Palmer CGS, and Woodward JA 2003. Detecting genotype combinations that increase risk for disease: The maternal-fetal genotype incompatibility test. *Genet Epidemiol* 24:1–13.
- Smits JMA, Claas FHJ, van Houwelingen HC, and Persijn GG 1998. Do noninherited maternal antigens (NIMA) enhance renal graft survival? *Transpl Int* 11:82–88.
- Soler JM and Blangero J 2003. Longitudinal familial analysis of blood pressure involving parametric (co)variance functions. *BMC Genetics* 4, no. Suppl 1:S87.
- Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, . . . , and Loos RJF 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42, no. 11:937–948.
- Spinka C, Carroll RJ, and Chatterjee N 2005. Analysis of case-control studies of genetic and environmental factors with missing genetic information and haplotype-phase ambiguity. *Genet Epidemiol* 29, no. 2:108–127.
- Sun W 2012. A statistical framework for eQTL mapping using RNA-seq data. *Biometrics* 68:1–11.
- Tavtigian SV, Simard J, Teng DH, Abtin V, Baumgard M, Beck A, Camp NJ, Carillo AR, Chen Y, Dayananth P, Desrochers M, Dumont M, Farnham JM, Frank D, Frye C, Ghaffari S, Gupte JS, Hu R, Iliev D, Janecki T, Kort EN, Laity KE, Leavitt A, Leblanc G, McArthur-Morrison J, Pederson A, Penn B, Peterson KT, Reid JE, Richards S, Schroeder M, Smith R, Snyder SC, Swedlund B, Swensen J, Thomas A, Tranchant M, Woodland AM, Labrie F, Skolnick MH, Neuhausen S, Rommens J, and Cannon-Albright LA 2001. A candidate prostate cancer susceptibility gene at chromosome 17p. *Nat Genet* 27, no. 2:172–180.



- Thiele TN 1899. Om iagttagelseslaere halvvarianter. Overs Vid Sels Forh pp. 135–141.
- Thompson EA 2008. The IBD process along four chromosomes. *Theor Popul Biol* 73, no. 3:369–373.
- Tibshirani R 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B (Methodological)* 58, no. 1:267–288.
- Tsonaka R, De Visser MCH, and Houwing-Duistermaat J 2013. Estimation of genetic effects in multiple cases family studies using penalized maximum likelihood methodology. *Biostatistics* 14:220–231.
- Tsonaka R, van der Helm-van Mil AHM, and Houwing-Duistermaat JJ 2012. A two-stage mixed-effects model approach for gene-set analyses in candidate gene studies. *Statist Med* 31, no. 11-12:1190–1202.
- Umbach DM and Weinberg CR 1997. Designing and analysing case-control studies to exploit independence of genotype and exposure. *Stat Med* 16:1731–1743.
- van der Woude D, Houwing-Duistermaat JJ, Toes REM, Huizinga TWJ, Thomson W, Worthington J, van der Helm-van Mil AHM, and de Vries RRP 2009. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. *Arthritis Rheum* 60:916–23.
- van der Woude D, Lie BA, Lundström E, Balsa A, Feitsma AL, Houwing-Duistermaat JJ, Verduijn W, Nordang GBN, Alfredsson L, Klareskog L, Pascual-Salcedo D, Gonzalez-Gay MA, Lopez-Nevot MA, Valero F, Roep BO, Huizinga TWJ, Kvien TK, Martín J, Padyukov L, de Vries RRP, and Toes REM 2010. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1\*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein . *Arthritis and rheumatism* 62:1236–45.
- Vollmar T, Maus B, Wurtz RP, Gillessen-Kaesbach G, Horsthemke B, Wiczorek D, and Boehringer S 2008. Impact of geometry and viewing angle on classification accuracy of 2d based analysis of dysmorphic faces. *European Journal of Medical Genetics* 51, no. 1:44–53.
- Wang WYS, Barratt BJ, Clayton DG, and Todd JA 2005. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6, no. 2:109–118.
- Weinberg, C R 1999. Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet* 65:229–35.
- Weir BS 1990. *Genetic Data Analysis: Methods for Discrete Population Genetic Data*. Sinauer Associates Incorporated.
- Winter RM 1996. What's in a face? *Nature Genetics* 12, no. 2:124–129.
- Witte JS 2010. Genome-wide association studies and beyond. *Annu Rev Public Health* 31:9–20.
- World Health Organization 2008. *Atlas: Multiple sclerosis resources in the world*. WHO - Atlas Project Neurology .

- Worthington J, Olliver WFR, Leach MK, Smith I, Hay EM, Thomson W, Pepper L, Carthy D, Farhan A, Martin S, Dyer P, Davison J, Bamber S, and Silman AJ 1994. Research practice the arthritis and rheumatism council's national repository of family material: Pedigrees from the first 100 rheumatoid arthritis families containing affected sibling pairs. *Br J Rheumatol* 33:970–976.
- Xing G and Xing C 2010. Adjusting for covariates in logistic regression models. *Genet Epidemiol* 34:769–771.
- Yousefi P, Huen K, Schall RA, Decker A, Elboudwarej E, Quach H, Barcellos L, and Holland N 2013. Considerations for normalization of DNA methylation data by illumina 450K BeadChip assay in population studies. *Epigenetics* 8:1141–1152.
- Zaitlen N, Lindström S, Pasaniuc B, Cornelis M, Genovese G, Pollack S, Barton A, Bickelböller H, Bowden DW, Eyre S, Freedman BI, Friedman DJ, Field JK, Groop L, Haugen A, Heinrich J, Henderson BE, Hicks PJ, Hocking LJ, Kolonel LN, Landi MT, Langefeld CD, Le Marchand L, Meister M, Morgan AW, Raji OY, Risch A, Rosenberger A, Scherf D, Steer S, Walshaw M, Waters KM, Wilson AG, Wordsworth P, Zienolddiny S, Tchetgen ET, Haiman C, Hunter DJ, Plenge RM, Worthington J, Christiani DC, Schaumberg DA, Chasman DI, Altshuler D, Voight B, Kraft P, Patterson N, and Price AL 2012a. Informed conditioning on clinical covariates increases power in case-control association studies. *PLoS Genet* 8:e1003032.
- Zaitlen N, Pasaniuc B, Patterson N, Pollack S, Voight B, Groop L, Altshuler D, Henderson BE, Kolonel LN, Marchand LL, Waters K, Haiman CA, Stranger BE, Dermitzakis ET, Kraft P, and Price AL 2012b. Analysis of case-control association studies with known risk variants. *Bioinformatics* 28:1729–1737.
- Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, and Ehm MG 2002. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 53, no. 2:79–91.
- Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, Craig DW, Redman M, Gershon ES, and Liu C 2010. Genetic control of individual differences in gene-specific methylation in human brain. *The American Journal of Human Genetics* 86:411–419.
- Zhao S, Cai T, and Li H 2014. More powerful genetic association testing via a new statistical framework for integrative genomics. To appear in *Biometrics* -.
- Zheng Y, Heagerty PJ, Hsu L, and Newcomb P 2010. On combining family-based and population-based case-control data in association studies. *Biometrics* 66:1024–33.

