ARIC Manuscript Proposal #2365

PC Reviewed: 5/13/14	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Efficient multi-locus association test for SNP set significance analysis

b. Abbreviated Title (Length 26 characters):

2. Writing Group:

Working group members:

Baolin Wu
James Pankow
Weihua Guan

Other interested investigators are welcome to join the writing group

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __BW__ **[please confirm with your initials electronically or in writing]**

First author: **Baolin Wu**
Address: Division of Biostatistics
University of Minnesota
A460 Mayo Bldg., MMC 303
420 Delaware St., S.E.
Minneapolis, MN 55455

Phone: 612-626-4765 Fax: 612-624-0660
E-mail: baolin@umn.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name:
Address:

Phone: Fax:
E-mail:

3. Timeline:
We anticipate a manuscript ready to submit for Publications Committee review in summer, 2014.

4. Rationale:

When analysing low frequency SNPs generated by next-generation sequencing, power is often a concern for single-SNP association tests. One common approach is to analyze a set of SNPs simultaneously, which aggregates information across multiple variants and reducing the burden of extreme multiplicity of testing (Liu et al., 2010; Madsen and Browning, 2009; Neale et al., 2011; Price et al., 2010; Wu et al., 2010, 2011; Wang and Abbott, 2008). One research question concerns the significance of a given SNP set, where we test whether the SNP set contains any risk variant associated with the phenotype. Existing methods are broadly based on the burden test (Li and Leal, 2008; Madsen and Browning, 2009) and variance component based score test approaches (SKAT; Wu et al., 2010, 2011). The burden test aggregates the variant scores within a set and performs well under strong assumptions of similar variant risks. The variance component based approach assumes unequal variant risks, and performs well in the presence of both protective and deleterious variants. The SKAT approach is based on the principle of the score test.

The score test has a considerable computational advantage since typically just one common null model needs to be estimated. However the score test may not have the best detection power especially for low-frequency and rare variants. We propose to develop alternative approaches based on the Wald and likelihood ratio test, which could offer much improved detection power compared to the score test (especially for association studies of discrete trait and rare variants).

5. Main Hypothesis/Study Questions:

We will develop statistical methods for testing whether a SNP set contains any risk variants. We will apply the proposed methods to the ARIC data to identify SNP sets associated with type 2 diabetes and related phenotypes.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

For single-SNP association test, we first assume a linear model. Let G denote the genotype score, Y the trait of interest, and Z the covariates to be adjusted for. For a set of K SNPs, the Y-G association can be jointly tested in the following regression model:

$$Y = Z\alpha + \sum_{j=1}^{K} G_j\beta_j + \varepsilon$$

where the null hypothesis: H_0: $\beta_1 = \cdots = \beta_K = 0$. To reduce the number of degrees of freedom in the test, the burden test approach typically assumes $\beta_1 = \cdots = \beta_K$. To allow

the effect β_j vary across SNPs with different directions, score-type tests were developed by Pan (2009) and Wu et al. (2010; 2011). Note that the score test statistic is derived under the null hypothesis. Under the alternative hypothesis that some SNPs in the set are associated with the trait, the variance estimated is (in general upward) biased, which may lead to loss of power. We consider other type of test statistic in this proposal. Specifically, we propose to use the t-statistics or likelihood ratio test (LRT) from the single-SNP linear regression model, and construct new test statistics. Under the null hypothesis, these single-SNP test statistics will approximately follow the multivariate normal distribution with zero mean and some covariance matrix, which can be readily computed based on the score statistics. We will demonstrate the power of proposed methods using extensive simulations.

For discrete traits, we will adopt a generalized linear model (GLM), such that Y independently follows an exponential family distribution. The LRT statistic for each SNP can be readily obtained from the single-SNP GLM, and signed square root of the LRT will be used to construct the new test statistics.

We will use the developed methods to answer research questions relevant to public health in the Atherosclerosis Risk in Communities (ARIC) study. Specifically, we will apply our methods to re-examine the association of exome chip SNPs with type 2 diabetes and related traits, including fasting glucose, fasting insulin, 2-hour glucose after an oral glucose tolerance test, and hemoglobin A1c in non-diabetic white samples in ARIC. In addition, we will compare results from the proposed methods to existing alternative methods.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes __x__ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No
(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ___x_ Yes ____ No Limited to ancestry information obtained from AIMs or GWAS markers

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"?
__x__ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this

proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

_____ Yes ___x____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

ARIC GWAS results for each of the glycemia traits have already been published as part of international meta-analyses conducted by the DIAGRAM Consortium (diabetes status in European populations) and MAGIC Consortium (diabetes-related continuous phenotypes in European populations):

Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segrè AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Benediktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bengtsson Boström K, Bravenboer B, Bumpstead S, Burtt NP, Charpentier G, Chines PS, Cornelis M, Couper DJ, Crawford G, Doney AS, Elliott KS, Elliott AL, Erdos MR, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnson PR, Jørgensen T, Kao WH, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lieverse A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JR, Petersen AK, Platou C, Proença C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shrader P, Sigurdsson G, Sparsø T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, Tichet J, Tuomi T, van Dam RM, van Haeften TW, van Herpt T, van Vliet-Ostaptchouk JV, Walters GB, Weedon MN, Wijmenga C, Witteman J, Bergman RN, Cauchi S, Collins FS, Gloyn AL, Gyllensten U, Hansen T, Hide WA, Hitman GA, Hofman A, Hunter DJ, Hveem K, Laakso M, Mohlke KL, Morris AD, Palmer CN, Pramstaller PP, Rudan I, Sijbrands E, Stein LD, Tuomilehto J, Uitterlinden A, Walker M, Wareham NJ, Watanabe RM, Abecasis GR, Boehm BO, Campbell H, Daly MJ, Hattersley AT, Hu FB, Meigs JB, Pankow JS, Pedersen O, Wichmann HE, Barroso I, Florez JC, Frayling TM, Groop L, Sladek R, Thorsteinsdottir U, Wilson JF, Illig T, Froguel P, van Duijn CM, Stefansson K, Altshuler D, Boehnke M, McCarthy MI; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010; 42 :579-89.

Rasmussen-Torvik LJ, Alonso A, Li M, Kao W, Kottgen A, Yan Y, Couper D, Boerwinkle E, Bielinski SJ, Pankow JS. Impact of repeated measures and sample selection on genome-wide association studies of fasting glucose. Genet Epidemiol 2010; 34: 665-673.

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T, Grarup N, Sparsø T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccasecca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen YD, Chines P, Clarke R, Coin LJ, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day IN, de Geus EJ, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CN, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AF, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurethsson G, Sijbrands EJ, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen AC, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JC, Yarnell JW, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium, Borecki IB, Loos RJ, Meneton P, Magnusson PK, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WH, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BW, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I.

New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nature Genetics 2010; 42: 105-116. PMCID: PMC3018764

Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, Mägi R, Strawbridge RJ, Rehnberg E, Gustafsson S, Kanoni S, Rasmussen-Torvik LJ, Yengo L, Lecoeur C, Shungin D, Sanna S, Sidore C, Johnson PC, Jukema JW, Johnson T, Mahajan A, Verweij N, Thorleifsson G, Hottenga JJ, Shah S, Smith AV, Sennblad B, Gieger C, Salo P, Perola M, Timpson NJ, Evans DM, Pourcain BS, Wu Y, Andrews JS, Hui J, Bielak LF, Zhao W, Horikoshi M, Navarro P, Isaacs A, O'Connell JR, Stirrups K, Vitart V, Hayward C, Esko T, Mihailov E, Fraser RM, Fall T, Voight BF, Raychaudhuri S, Chen H, Lindgren CM, Morris AP, Rayner NW, Robertson N, Rybin D, Liu CT, Beckmann JS, Willems SM, Chines PS, Jackson AU, Kang HM, Stringham HM, Song K, Tanaka T, Peden JF, Goel A, Hicks AA, An P, Müller-Nurasyid M, Franco-Cereceda A, Folkersen L, Marullo L, Jansen H, Oldehinkel AJ, Bruinenberg M, Pankow JS, North KE, Forouhi NG, Loos RJ, Edkins S, Varga TV, Hallmans G, Oksa H, Antonella M, Nagaraja R, Trompet S, Ford I, Bakker SJ, Kong A, Kumari M, Gigante B, Herder C, Munroe PB, Caulfield M, Antti J, Mangino M, Small K, Miljkovic I, Liu Y, Atalay M, Kiess W, James AL, Rivadeneira F, Uitterlinden AG, Palmer CN, Doney AS, Willemsen G, Smit JH, Campbell S, Polasek O, Bonnycastle LL, Hercberg S, Dimitriou M, Bolton JL, Fowkes GR, Kovacs P, Lindström J, Zemunik T, Bandinelli S, Wild SH, Basart HV, Rathmann W, Grallert H; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium, Maerz W, Kleber ME, Boehm BO, Peters A, Pramstaller PP, Province MA, Borecki IB, Hastie ND, Rudan I, Campbell H, Watkins H, Farrall M, Stumvoll M, Ferrucci L, Waterworth DM, Bergman RN, Collins FS, Tuomilehto J, Watanabe RM, de Geus EJ, Penninx BW, Hofman A, Oostra BA, Psaty BM, Vollenweider P, Wilson JF, Wright AF, Hovingh GK, Metspalu A, Uusitupa M, Magnusson PK, Kyvik KO, Kaprio J, Price JF, Dedoussis GV, Deloukas P, Meneton P, Lind L, Boehnke M, Shuldiner AR, van Duijn CM, Morris AD, Toenjes A, Peyser PA, Beilby JP, Körner A, Kuusisto J, Laakso M, Bornstein SR, Schwarz PE, Lakka TA, Rauramaa R, Adair LS, Smith GD, Spector TD, Illig T, de Faire U, Hamsten A, Gudnason V, Kivimaki M, Hingorani A, Keinanen-Kiukaanniemi SM, Saaristo TE, Boomsma DI, Stefansson K, van der Harst P, Dupuis J, Pedersen NL, Sattar N, Harris TB, Cucca F, Ripatti S, Salomaa V, Mohlke KL, Balkau B, Froguel P, Pouta A, Jarvelin MR, Wareham NJ, Bouatia-Naji N, McCarthy MI, Franks PW, Meigs JB, Teslovich TM, Florez JC, Langenberg C, Ingelsson E, Prokopenko I, Barroso I. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nature Genetics 2012; 44: 991-1005. PMCID: PMC3433394

Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, Vollenweider P, Lyssenko V, Bouatia-Naji N, Dupuis J, Jackson AU, Kao WH, Li M, Glazer NL, Manning AK, Luan J, Stringham HM, Prokopenko I, Johnson T, Grarup N, Boesgaard TW, Lecoeur C, Shrader P, O'Connell J, Ingelsson E, Couper DJ, Rice K, Song K, Andreasen CH, Dina C, Köttgen A, Le Bacquer O, Pattou F, Taneera J, Steinthorsdottir V, Rybin D, Ardlie K, Sampson M, Qi L, van Hoek M, Weedon MN, Aulchenko YS, Voight BF, Grallert H, Balkau B, Bergman RN, Bielinski SJ, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Buchanan TA, Bumpstead SJ, Cavalcanti-Proença C, Charpentier

G, Chen YD, Chines PS, Collins FS, Cornelis M, J Crawford G, Delplanque J, Doney A, Egan JM, Erdos MR, Firmann M, Forouhi NG, Fox CS, Goodarzi MO, Graessler J, Hingorani A, Isomaa B, Jørgensen T, Kivimaki M, Kovacs P, Krohn K, Kumari M, Lauritzen T, Lévy-Marchal C, Mayor V, McAteer JB, Meyre D, Mitchell BD, Mohlke KL, Morken MA, Narisu N, Palmer CN, Pakyz R, Pascoe L, Payne F, Pearson D, Rathmann W, Sandbaek A, Sayer AA, Scott LJ, Sharp SJ, Sijbrands E, Singleton A, Siscovick DS, Smith NL, Sparsø T, Swift AJ, Syddall H, Thorleifsson G, Tönjes A, Tuomi T, Tuomilehto J, Valle TT, Waeber G, Walley A, Waterworth DM, Zeggini E, Zhao JH; GIANT consortium; MAGIC investigators, Illig T, Wichmann HE, Wilson JF, van Duijn C, Hu FB, Morris AD, Frayling TM, Hattersley AT, Thorsteinsdottir U, Stefansson K, Nilsson P, Syvänen AC, Shuldiner AR, Walker M, Bornstein SR, Schwarz P, Williams GH, Nathan DM, Kuusisto J, Laakso M, Cooper C, Marmot M, Ferrucci L, Mooser V, Stumvoll M, Loos RJ, Altshuler D, Psaty BM, Rotter JI, Boerwinkle E, Hansen T, Pedersen O, Florez JC, McCarthy MI, Boehnke M, Barroso I, Sladek R, Froguel P, Meigs JB, Groop L, Wareham NJ, Watanabe RM. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nature Genetics 2010; 42: 142-148. PMCID: PMC2922003

Soranzo N, Sanna S, Wheeler E, Gieger C, Radke D, Dupuis J, Bouatia-Naji N, Langenberg C, Prokopenko I, Stolerman E, Sandhu MS, Heeney MM, Devaney JM, Reilly MP, Ricketts SL, Stewart AF, Voight BF, Willenborg C, Wright B, Altshuler D, Arking D, Balkau B, Barnes D, Boerwinkle E, Böhm B, Bonnefond A, Bonnycastle LL, Boomsma DI, Bornstein SR, Böttcher Y, Bumpstead S, Burnett-Miller MS, Campbell H, Cao A, Chambers J, Clark R, Collins FS, Coresh J, de Geus EJ, Dei M, Deloukas P, Döring A, Egan JM, Elosua R, Ferrucci L, Forouhi N, Fox CS, Franklin C, Franzosi MG, Gallina S, Goel A, Graessler J, Grallert H, Greinacher A, Hadley D, Hall A, Hamsten A, Hayward C, Heath S, Herder C, Homuth G, Hottenga JJ, Hunter-Merrill R, Illig T, Jackson AU, Jula A, Kleber M, Knouff CW, Kong A, Kooner J, Köttgen A, Kovacs P, Krohn K, Kühnel B, Kuusisto J, Laakso M, Lathrop M, Lecoeur C, Li M, Li M, Loos RJ, Luan J, Lyssenko V, Mägi R, Magnusson PK, Mälarstig A, Mangino M, Martínez-Larrad MT, März W, McArdle WL, McPherson R, Meisinger C, Meitinger T, Melander O, Mohlke KL, Mooser VE, Morken MA, Narisu N, Nathan DM, Nauck M, O'Donnell C, Oexle K, Olla N, Pankow JS, Payne F, Peden JF, Pedersen NL, Peltonen L, Perola M, Polasek O, Porcu E, Rader DJ, Rathmann W, Ripatti S, Rocheleau G, Roden M, Rudan I, Salomaa V, Saxena R, Schlessinger D, Schunkert H, Schwarz P, Seedorf U, Selvin E, Serrano-Ríos M, Shrader P, Silveira A, Siscovick D, Song K, Spector TD, Stefansson K, Steinthorsdottir V, Strachan DP, Strawbridge R, Stumvoll M, Surakka I, Swift AJ, Tanaka T, Teumer A, Thorleifsson G, Thorsteinsdottir U, Tönjes A, Usala G, Vitart V, Völzke H, Wallaschofski H, Waterworth DM, Watkins H, Wichmann HE, Wild SH, Willemsen G, Williams GH, Wilson JF, Winkelmann J, Wright AF; WTCCC, Zabena C, Zhao JH, Epstein SE, Erdmann J, Hakonarson HH, Kathiresan S, Khaw KT, Roberts R, Samani NJ, Fleming MD, Sladek R, Abecasis G, Boehnke M, Froguel P, Groop L, McCarthy MI, Kao WH, Florez JC, Uda M, Wareham NJ, Barroso I, Meigs JB. Common variants at ten genomic loci influence hemoglobin A1C levels via glycemic and non-glycemic pathways. Diabetes 2010; 59: 3229-3239. PMCID: PMC2992787

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __x__ Yes ____ No

11.b. If yes, is the proposal

_x__ A. primarily the result of an ancillary study (list number* 2013.19_**)**

___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)

2007.02 (CARe, genotyping in African Americans)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PUBMED Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to Pubmed central.

References:

Li,B. and Leal,S.M. (2008) Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. The American Journal of Human Genetics, 83 (3), 311–321.

Liu,J.Z., Mcrae,A.F., Nyholt,D.R., Medland,S.E., Wray,N.R., Brown,K.M., Investigators,A., Hayward,N.K., Montgomery, G.W., Visscher,P.M., Martin,N.G. and Macgregor,S. (2010) A versatile gene-based test for genomewide association studies. The American Journal of Human Genetics, 87 (1), 139–145.

Madsen,B.E. and Browning,S.R. (2009) A groupwise association test for rare mutations using a weighted sum statistic. PLOS Genetics, 5 (2), e1000384.

Neale,B.M., Rivas,M.A., Voight,B.F., Altshuler,D., Devlin,B., Orho-Melander,M., Kathiresan,S., Purcell,S.M., Roeder,K. and Daly,M.J. (2011) Testing for an unusual distribution of rare variants. PLoS Genet, 7 (3), e1001322.

Pan, W., Asymptotic tests of association with multiple SNPs in linkage disequilibrium. Genetic Epidemiology, 2009. 33(6): p. 497-507.

Wang,K. and Abbott,D. (2008) A principal components regression approach to multilocus genetic association studies. Genetic Epidemiology, 32 (2), 108–118.

Wu,M., Lee,S., Cai,T., Li,Y., Boehnke,M. and Lin,X. (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. The American Journal of Human Genetics, 89 (1), 82–93.

Wu,M.C., Kraft,P., Epstein,M.P., Taylor,D.M., Chanock,S.J., Hunter,D.J. and Lin,X. (2010) Powerful SNP-Set analysis for case-control genome-wide association studies. The American Journal of Human Genetics, 86 (6),
929–942.