S T R E G A

Strengthening the **Re**porting of **G**enetic **A**ssociations

Report on Workshop held June 15-16, 2006
Ottawa, Ontario, Canada, and subsequent work

January 31, 2008

GenomeCanada

Public Health Agency of Canada
Agence de santé publique du Canada

Report on Workshop held June 15-16, 2006 Ottawa, Ontario, Canada, and subsequent work

Table of Contents

Report on workshop held 15-16 June 2006 Ottawa, Ontario, Canada, and subsequent work

In the "Needs, gaps, and opportunities analysis for enhancing the functionality of the Human Genome Epidemiology Network," the most pressing need identified was for improvement in the reporting of studies of genetic associations and related interactions. We now report on a multidisciplinary workshop organised in response to this need in Ottawa in June 2006 and subsequent work.

1. Rationale for strengthening the reporting of genetic associations

The rapidly evolving evidence on genetic associations is crucial to integrating human genomics into the practice of medicine and public health (1, 2). Genetic factors are likely to have an impact on the occurrence of numerous common diseases, and therefore identifying and characterizing the associated risk, or protection, will be important in improving understanding of etiology and potentially for developing interventions that might be based on genetic information.

The number of publications on gene-disease associations has increased tremendously, with the number each year having more than doubled between 2001 and 2007, with more than 30,000 published articles during that time (3, 4). Articles on genetic associations have been published in about 1500 journals, in several languages.

Although there are a number of similarities between genetic association studies and "classical" observational epidemiologic studies of lifestyle and environmental factors, the former present several specific challenges including an unprecedented volume of new data (5, 6) and the likelihood of very small individual effects. Genes may operate in complex pathways with gene-environment and gene-gene interactions (7). Moreover, the current evidence base on gene-disease associations is fraught with methodological problems (8-10). These include inadequate statistical power; flawed study design; suboptimal study conduct and biased analyses; lack of standardization among studies; selective reporting of "positive" results; and poor or incomplete reporting of results even from well-conducted studies (11).

The adequate reporting of studies of the association between diseases and putative risk factors is of importance in

- assembling empirical evidence regarding methodological biases which might affect this type of study, and thereby help improve study design and conduct in the longer term;
- minimising the potential problems of selective reporting and publication bias; and
- facilitating the synthesis of knowledge.

Although several commentaries on the conduct and/or appraisal of genetic association studies have been published that cover issues in reporting (12-39), their recommendations differ. For example, some papers suggest that replication of findings should be part of any publication (12, 13, 16, 17, 23, 26, 34-36) while others consider this suggestion unnecessary or even unreasonable, such as when a novel hypothesis is tested in a large

well-conducted study (21, 40-44). In many, the guidance has focused on conduct of genetic studies rather than reporting (13-15, 17, 19, 20, 22, 23, 25, 30-32, 35, 36) or has focused on association studies for specific diseases (14, 15, 17, 19, 22, 23, 25, 26, 31-38).

Despite increasing recognition of these problems, the quality of reporting genetic association studies is not optimal (45-49). For example, an assessment of a random sample of 315 genetic association studies published from 2001 to 2003 found that most studies provided some qualitative descriptions of the study participants (e.g. origin and enrolment criteria), but reporting of quantitative descriptors such as age and gender was variable, as was reporting of methods that allow readers to assess potential biases (e.g., number of exclusions or number of samples that could not be genotyped) (49). Only a minority of studies described methods to validate genotyping or mentioned whether research staff involved in this were blinded to outcome. The same problems persisted in a smaller sample of studies published in 2006 (49).

Lack of transparency and incomplete reporting have raised concerns in a range of health research fields (11, 50-53) and poor reporting has been associated with biased estimates of effects (54). To help remedy this problem some groups have developed evidence-based reporting guidelines. For example, the Consolidated Standards of Reporting Trials (CONSORT) Statement (55-57), which provides a 22-item checklist and flow diagram, has been found to improve the reporting of randomized clinical trials (RCTs) (57). Using a similar template, the epidemiology community has recently developed a reporting

guidance for cross-sectional, case-control, and cohort studies, the **ST**rengthening the **R**eporting of **OB**servational studies in **E**pidemiology (STROBE) Statement (58-64).

We therefore organised a workshop on strengthening the reporting of genetic association studies that would build on the experience of the Human Genome Epidemiology Network (HuGENet), and on the work done by the STROBE Working Group. HuGENet is a global initiative committed to the development and integration of the knowledge base on human genetic variants and health (http://www.cdc.gov/genomics/hugenet); however, while benefiting from the experience of this network, we sought to develop guidance that would have the broadest applicability possible. The STROBE reporting guidance is a strong basis because it is evidence-based and involved extensive consultations in the epidemiological research community (58, 59, 61, 62, 64-69). The unique contribution of the proposed workshop was the focus on gene-disease association studies.

The workshop and subsequent work was intended to promote clear reporting of genetic association studies. Clear reporting will help journal editors and researchers identify methodological biases in such studies and facilitate synthesis of knowledge about the role of genetic variation in population health. Identification of methodological biases will be crucial in design and interpretation of future genetic association studies. Therefore the workshop, funded by the Public Health Agency of Canada, the CIHR Institutes of Genetics (IG) and Nutrition, Metabolism and Diabetes (INMD), Genome Canada, Affymetrix, DNA Genotek and TrialStat, and with in-kind support from GeneSens, aimed

to bring together an international multidisciplinary group of researchers and journal editors to initiate and implement the development of guidance.

Objectives

The objectives of the workshop and subsequent work were:

- To provide guidance for reporting the results of studies of gene-disease associations that (1) is evidence based, (2) represents a consensus of epidemiologists, geneticists and statisticians, and (3) is built on the STROBE statement.
- To identify gaps and areas of controversy in the evidence regarding potential biases in genetic association studies. This will guide research that will inform future revisions to the guidance.

We report on the workshop and subsequent work in relation to these objectives.

2. Development of guidance for reporting the results of genetic association studies

2.1 Workshop preparation

The planning of the workshop was taken forward by a co-coordinating group of seven people (Julian Little, Julian Higgins, John Ioannidis, David Moher, France Gagnon, Eric von Elm, Muin Khoury). Prior to the workshop, an electronic search was performed to identify existing guidance on the reporting of genetic association studies. Workshop participants were asked to identify any additional guidance. Several workshop participants prepared brief presentations on existing reporting guidance, empirical evidence on reporting of genetic association studies, the development of the STROBE Statement, and on several key areas for discussion identified on the basis of consultations prior to the workshop. These included the selection and participation of study subjects,

rationale for choice of genes and variants investigated, genotyping errors, methods for inferring haplotypes, population stratification, assessment of Hardy-Weinberg equilibrium, multiple testing, reporting of quantitative (continuous) outcomes, selectively reporting study results, joint effects and inference of causation in single studies. Additional resources to inform workshop participants were the HuGENet handbook (70, 71), examples of data extraction forms from systematic reviews/meta-analyses, papers on guideline development (55, 56) and the checklists developed for STROBE.

2.2 Selection of workshop participants

Seventy-four people were invited to participate in the workshop including epidemiologists, geneticists, statisticians, journal editors and graduate students, including contributors to and coordinators of the STROBE Initiative. Thirty three people participated (list in Appendix A).

In broad terms, epidemiologists, geneticists and statisticians were invited to the workshop based on their involvement in one or more of

- development of STROBE statement
- development of genetic association databases or
- publications on methodological issues related to genetic association studies.

We refer to broad principles as it would not have been feasible to invite all investigators who would qualify for at least one criterion. In this group, 18 participated, 15 responded that they would have liked to participate but were unable to do so, and 3 did not respond.

Four graduate students in epidemiology from the University of Ottawa participated in the workshop, acted as rapporteurs, and contributed to the development of a manuscript.

Again in broad terms, the journal editors were selected on the basis of (a) general influence of journal; (b) ranking in terms of number of genetic association articles published; (c) previous publication of guidance; (d) connection with STROBE; (e) connection with HuGENet. Nine journal editors took part in the workshop – the journals represented were:

- *American Journal of Human Genetics*
- *Epidemiology*
- *European Journal of Epidemiology*
- *Genetics in Medicine*
- *International Journal of Epidemiology*
- *Lancet*
- *Nature Genetics*
- *Paediatric and Perinatal Epidemiology*
- *PLoS* (Public Library of Science).

A further seven responded that they would have liked to participate but were unable to do so. These editors were from

- *American Journal of Epidemiology*
- *Arthritis & Rheumatism*
- *Cancer Epidemiology Biomarkers & Prevention*

- *Genetic Epidemiology*
- *Human Genetics*
- *Journal of Clinical Epidemiology*
- *Nature Reviews Genetics*
- *PLoS Med.*

No response was received from the editors of 13 journals (*American Journal of Medical Genetics; American Journal of Respiratory and Critical Care Medicine; Birth Defects Research Part A (Clinical and Molecular Teratology); Diabetes; European Journal of Human Genetics; Human Molecular Genetics; International Journal of Cancer; Journal of the American Medical Association; Metabolism; New England Journal of Medicine; PLoS Genetics; Schizophrenia Research; Science).*

In the course of teaching a session in the WHO/PAHO Evidence-based Medicine and Health Technology Assessment Training Course for health policy decision-makers from the People's Republic of China in Ottawa in March 2006, Julian Little met a representative of the Ministry of Health from that country. As there are a large number of publications on genetic association studies in Chinese journals (see Appendix B), following that meeting, a formal letter translated into Mandarin was sent to this representative inviting a representative of one of these journals to participate in the workshop. Unfortunately, no response was received.

2.3 STROBE

The STROBE initiative was established in 2004, starting with a workshop and a dedicated website (www.strobe-statement.org). Following a systematic search for publications on reporting of observational studies, a central archive of published recommendations, guidelines and checklists relating to the reporting of epidemiological studies, quality assessment tools, empirical studies of reporting and other methodological research articles was established. The STROBE group decided at an early stage to focus on the three study designs that are most widely used in analytical observational epidemiology: cohort studies, case-control studies, and cross-sectional studies. Twenty-three epidemiologists, methodologists, statisticians, researchers conducting observational studies, and editors of general medicine journals and specialist epidemiology and public health journals participated in a two-day workshop in October 2004. During the workshop, three working groups identified items deemed important to include in checklists for each of the three study designs, based on a provisional list of items extracted from the literature. Wherever possible, items were revised in order to make them applicable to all three study designs. After the workshop, the participants, as well as additional scientists and editors were invited to comment on the draft checklist. The checklist subsequently underwent further revisions, which were made available on the website with a summary of received comments and a renewed invitation to comment. The STROBE Statement was published in October-November 2007 in seven journals (Ann.Intern.Med.; BMJ, Bull.World Health Organ.; Epidemiology; Lancet; PLoS Med.; Prev.Med.) (58-64), and an explanation and elaboration in three (Ann.Intern.Med.; Epidemiology; PLoS Med.) (72-74).

Of note is the emphasis on strengthening reporting as distinct from developing reporting standards (67, 68), and as distinct from focusing on how research should be done as this might stifle methodological innovation (75). Since the publication of the STROBE Statement, it has been suggested that future versions of the STROBE guidance should include the consideration of incubation periods for risk factors and diseases, biological plausibility, and clear definition and presentation of results on host factors (76). In commentaries about STROBE, it is interesting that one of the authors emphasized that these reporting guidelines do not constitute an instrument to evaluate the quality of research (68) and it has been suggested that the authors of STROBE should expressly discourage the use of the guideline for the evaluation of studies or study results, and that "the blindly applied rule" should not "trump the creative exception" (67).

To harmonize our guidance for genetic association studies with more general guidelines for observational epidemiologic studies, we communicated with the STROBE group during the development process and sought their comments on STREGA. We also provided comments on the developing STROBE statement and its associated explanation and elaboration document (the contributions of individuals in the STREGA initiative is acknowledged in the STROBE papers).

2.4 Content of workshop and subsequent work

The two-day workshop was a mixture of plenary group sessions and breakout sessions. The plenary sessions included presentations on the key areas for discussion that had been

identified before the meeting, including the processes by which the CONSORT, STROBE and related reporting guides were developed, and previously proposed guidelines on the reporting of genetic association studies (a copy of the Agenda is included as Appendix C). Three breakout groups were tasked with considering the extent to which the draft STROBE guideline could be applied to genetic association studies of

- case-control
- cohort and
- cross-sectional designs.

Rapporteurs recorded the discussion. The meeting concluded with a moderated discussion to draw together the conclusions of the groups regarding what guidance could be offered now and the remaining gaps.

Further development of the STREGA guidance was achieved by several iterations of electronic correspondence among participants following the workshop over a period of 18 months (June 2006-January 2008), together with telephone discussions by members of the coordinating group.

One participant has declined to be included in the authorship, or to be acknowledged, on the grounds of (1) having become “more skeptical of the value of these “superstar” articles.” In particular, this participant was an author in the paper by the NCI-NHGRI Working Group on Replication in Association Studies (39), but now “question(s) its claims or even whether it faced the issues head-on”; (2) the field being in flux about major issues such as population stratification and the importance of Hardy Weinberg

equilibrium when good quality control measures are in place. Another participant withdrew, but on the grounds of feeling he ought to distance himself as a journal editor, and he was happy to be included in the acknowledgements. Two participants were present as observers and hence have not been included in the authorship.

2.5 The STREGA extension to the STROBE checklist

The STREGA extension to the STROBE checklist includes ten new items (one relating to reporting of variables, one to data sources/measurement, one to bias, four to statistical methods, one to main results, and two to other analyses) and modifies six others (Table 1).

2.6 Rationale for new and modified items in STREGA extension to STROBE

The rationale for the new and modified items is summarized in Table 2. We now comment on the main areas identified as of special interest in genetic association studies: genotyping error, population stratification, modelling haplotype variation, Hardy-Weinberg equilibrium and replication.

Genotyping errors

Genotyping errors can occur as a result of effects of the DNA sequence flanking the marker of interest, poor quality or quantity of the DNA extracted from biological samples, biochemical artifacts, poor equipment precision or equipment failure, or human error in sample handling, the conduct of the array or handling the data obtained from the array (77). In a commentary published in 2005 on the possible causes and consequences

of genotyping errors, it was observed that an increasing number of researchers were aware of the problem, but the effects of such errors had largely been neglected (77). The extent of genotyping errors has been reported to vary between about 1% and 30% (77-80), but is thought to be lower. In high-throughput centres where an error rate of 0.5% per genotype, for blind duplicates run on the same gel, has been observed (80). This lower error rate reflects an explicit choice of markers for which genotyping rates have been found to be highly repeatable and whose individual polymerase chain reactions (PCR) have been optimized. Hence we suggest specifying whether the genotyping was done in a high-throughput centre. Non-differential genotyping errors, i.e. those which do not differ systematically according to outcome status, will usually bias associations towards the null (81, 82), just as for other non-differential errors. The most marked bias occurs when genotyping sensitivity is poor and genotype prevalence is high (>85%) or, as the corollary, when genotyping specificity is poor and genotype prevalence is low (<15%) (81). When measurement of the environmental exposure has substantial error, genotyping errors of the order of 3% can lead to substantial under-estimation of the magnitude of an interaction effect (83). When there are systematic differences in genotyping according to outcome status (differential error), bias in any direction may occur. Unblinded assessment may lead to differential misclassification. For genome-wide association studies of single nucleotide polymorphisms (SNPs), differential misclassification between cases and controls can occur because of differences in DNA storage, collection or processing protocols, even when the genotyping itself meets the highest possible standards. (84). In this situation, using samples blinded to case-control status to determine the parameters for allele calling could still lead to differential

misclassification. To minimize such differential misclassification, it would be necessary to calibrate the software separately for each group. Hence, for case-control studies, we recommend that investigators report on whether or not genotyping was done blind to case-control status, and the reason for this decision.

Population stratification

Population stratification is the presence within a population of subgroups among which allele (or genotype; or haplotype) frequencies and disease risks differ. When the groups compared in the study differ in their proportions of the population subgroups, an association between the genotype and the disease being investigated may reflect the genotype being an indicator identifying population subgroup rather than a causal variant. In this situation, population subgroup is a confounder as it is associated with both genotype frequency and disease risk. There has been debate about the potential implications of population stratification for the validity of genetic association studies (85-99). Modelling the possible effect of population stratification (when no effort has been made to address it) suggests that the effect is likely to be small in most situations (91, 92, 94-96). Meta-analyses of 43 gene-disease associations comprising 697 individual studies show consistent associations across groups of different ethnic origin (96), and so provide evidence against a large effect of population stratification, hidden or otherwise. However, as studies of association and interaction typically address moderate or small effects and hence require large sample sizes, a small bias arising from population stratification, may be important (97). Study design (case-family control studies) and statistical methods (100) have been proposed to address population stratification, but so far few studies have

used these (49). Most of the early genome-wide association studies used either family-based designs or methods such as genomic control and principal components analysis (101, 102) to control for stratification. These approaches are of more concern for excluding bias when the identified genetic effects are very small (odds ratio <1.20). We recommend reporting of the methods used to address this potential problem, or stating that none was used, in order to enable empirical evidence to accrue.

Modelling haplotype variation

There has been increasing interest in modelling haplotype variation within candidate genes. Typically, the number of haplotypes observed within a gene is much smaller than the theoretical number of all possible haplotypes (103, 104). Motivation for utilizing haplotypes comes, in large part, from the fact that multiple SNPs may "tag" an untyped variant more effectively than a single typed variant. The subset of SNPs used in such an approach is referred to as "haplotype tagging" SNPs. Implicitly, an aim of haplotype tagging is to reduce the number of SNPs that have to be genotyped, while maintaining statistical power to detect an association with the phenotype.

In most current large-scale genetic association studies, data are collected as unphased multilocus genotypes (i.e. it is unknown which alleles are aligned together on particular segments of chromosome). It is common in such studies to use statistical methods to estimate haplotypes, but their accuracy and efficiency is debated (105-109). Some methods attempt to make use of a concept referred to as haplotype "blocks," but the results of these methods are sensitive to the specific definitions of the "blocks" (110,

111). Reporting of the methods used to infer haplotypes and the associated uncertainty should enhance our understanding of the possible effects of different methods of modelling haplotype variation on study results as well as enabling comparison and syntheses of results from different studies.

Information on common patterns of genetic variation revealed by the International Haplotype Map (HapMap) project (104) can be applied in the analysis of genome-wide association studies to infer genotypic variation at markers not typed directly in these studies (112-114). Essentially, these methods perform haplotype-based tests but make use of information on variation in a set of reference samples (e.g. HapMap) to guide the specific tests of association, collapsing a potentially large number of haplotypes into two classes (the allelic variation) at each marker. It is expected that these techniques will increase power in individual studies, and will aid in combining data across studies, and even across differing genotyping platforms. We recommend that it is made clear when such methods are used, and that they are clearly specified.

Hardy-Weinberg equilibrium

Hardy-Weinberg equilibrium (HWE) has become widely accepted as an underlying model in population genetics after Hardy (115) proposed the concept that allele frequencies at a genetic locus are stable within one generation of random mating; the assumption of HWE is equivalent to the independence of two alleles at a locus. Views differ as to whether testing for departure from HWE is a useful method of detecting errors or peculiarities in the data. In particular, it has been suggested that deviation from

HWE may be a sign of genotyping error (116-118). However, the statistical power to detect such errors by testing for departure from HWE is low and, in hypothetical data, the presence of HWE was not generally altered by the introduction of genotyping error (119). Furthermore, the assumptions underlying HWE, including random mating, lack of selection according to genotype, and absence of mutation or gene flow, are rarely met in human populations (120, 121). In five out of 42 gene-disease associations assessed in meta-analyses of almost 600 studies, the results of studies in which HWE was violated gave significantly different results from HWE-conforming studies (122). Moreover, the study suggested that exclusion of HWE-violating studies may result in loss of the statistical significance of some postulated gene-disease associations and that adjustment for the magnitude of deviation from HWE may also have the same consequence for some other gene-disease associations. We recommend that authors state whether HWE was considered, describing statistical tests or measures used, or any procedure used to allow for deviation from HWE (119).

Replication

Publications that present and synthesize data from several studies in a single report are becoming more common. In particular, many genome-wide association analyses describe several different study populations, sometimes with different study designs and genotyping platforms, and in various stages of discovery and replication (101, 123-144). In this setting, each of the constituent studies and the composite results should be fully described according to the STREGA reporting guideline. Although describing the

methods and results in sufficient detail would require substantial space, online options for depositing additional information on the study make this possible.

2.7 Issues considered already adequately covered by STROBE

Issues that were discussed during the workshop that were considered to be already adequately covered by STROBE are summarized in Table 3.

3. Gaps and areas of controversy in the evidence regarding potential biases in genetic association studies

In general, empirical evidence regarding the effects of study design, process and analysis on the results of genetic association studies is insufficient; transparency of reporting is thus essential for developing a better evidence base (Table 2). Transparent reporting will help address gaps in empirical evidence (45), such as the effects of incomplete participation and genotyping error. It will also help assess the impact of currently controversial issues such as population stratification, methods of inferring haplotypes, departure from Hardy-Weinberg equilibrium and multiple testing on effect estimates under different study conditions.

We made an explicit decision to focus attention on cross-sectional, case-control, and cohort studies, as had been done in STROBE. There is a need and opportunity to cover more specialized designs such as case-parent trio studies, other studies of cases and their relatives, and the case-only design. Subsequent to the workshop, there was a substantial amount of discussion about genome-wide association studies. Most of the issues related to conduct, which was not within the remit of the workshop, rather than reporting.

However, as more empirical evidence from these studies accrues, reporting of these studies would also merit further consideration.

4. Dissemination and evaluation

A manuscript reporting the process and outcome of the workshop and subsequent work has been prepared. Following the STROBE model, and also the (classical) model of the CONSORT Statement, we believe that it would be helpful to publish the STREGA Reporting Guidance in a number of journals simultaneously. Therefore we are sending the manuscript to the journal editors who participated in, or expressed an interest in participating in, the workshop, to seek their advice as to the most appropriate publication strategy. We believe that the range of investigators involved in genetic association research is so broad that it would be unlikely that a single journal publication would penetrate the community adequately.

In the manuscript, we invite journals to endorse STREGA, for example by updating their Instructions to Authors to include STREGA (and its URL), advising peer reviewers to use the checklist as a guide, and sending the checklist to authors with the reviewers' comments when a revision is requested. Thus, we consider that the STREGA guidance is a tool that can be used by authors, peer reviewers and editors to improve reporting.

After peer review, the STREGA guidance will also be posted on www.strega-statement.org with links to the HuGENet, STROBE and P³G websites. Comments will be used to refine future versions of the guidance.

We plan to evaluate the impact of the guidance by carrying out an empirical study comparing quality of reporting before and after its dissemination, as has recently been done for CONSORT (57). We note that an uncontrolled before-after comparison is a weak design, but it is the only quasi-experimental study that could be done easily. We will consider the possibilities of conducting an interrupted time series study or identifying case controlled before-after situations. We hope that the guidance will stimulate transparent and improved reporting of genetic association studies. In turn, better reporting of original studies would enable the further development of the empirical studies that are needed to support reporting guidelines such as STREGA.

Table 1. STREGA guidance, extended from STROBE Statement (58-64)

	Item number		*Extension for genetic association studies*
TITLE and ABSTRACT	1	(a) Indicate the study's design with a commonly used term in the title or the abstract.	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found.	
INTRODUCTION			
Background rationale	2	Explain the scientific background and rationale for the investigation being reported.	
Objectives	3	State specific objectives, including any pre-specified hypotheses.	***State if the study is the first report of a genetic association, a replication effort, or both.***
METHODS			
Study design	4	Present key elements of study design early in the paper.	

	Item number		*Extension for genetic association studies*
Setting	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up, and data collection.	
Participants	6	(a) **Cohort study –** Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. **Case-control study –** Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. **Cross-sectional study –** Give the eligibility criteria, and the sources and methods of selection of participants.	***Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant***.
		(b) **Cohort study –** For matched studies, give matching criteria and number of exposed and unexposed. **Case-control study –** For matched studies, give matching criteria and the number of controls per case.	
Variables	7	*(a)* Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	***(b) Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be associated with population***

	Item number		*Extension for genetic association studies*
			stratification (confounding by ethnic origin).
Data sources measurement	8*****	*(a)* For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	***(b) Describe specific laboratory methods, including source and storage of DNA, genotyping methods and platforms, error rates and call rates. State the laboratory/centre where genotyping was done****. Describe comparability of assessment methods if there is more than one group.*
Bias	9	*(a)* Describe any efforts to address potential sources of bias.	***(b) For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias, and explain what approach was used to deal with this.***

	Item number		*Extension for genetic association studies*
Study size	10	Explain how the study size was arrived at.	
Quantitative variables	11	*(a)* Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.	
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding.	***State software version used and options (or settings) chosen.***
		(b) Describe any methods used to examine subgroups and interactions.	
		(c) Explain how missing data were addressed.	
		(c) **Cohort study –** If applicable, explain how loss to follow-up was addressed. **Case-control study –** If applicable, explain how matching of cases and controls was addressed. **Cross-sectional study –** If applicable, describe analytical methods taking account of sampling strategy.	
		(e) Describe any sensitivity analyses.	

Item number			*Extension for genetic association studies*
			(f) State whether Hardy-Weinberg equilibrium was considered and, if so, how.
			(g) Describe any methods used for inferring genotypes or haplotypes.
			(h) Describe any methods used to assess or address population stratification.
			(i) Describe any methods used to address multiple tests or to control risk of false positive findings.
RESULTS			
Participants	13*****	*(a)* Report the numbers of individuals at each stage of the study – e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed.	***Include numbers in whom genotyping was attempted and numbers in whom genotyping was successful.***

	Item number		*Extension for genetic association studies*
		(b) Give reasons for non-participation at each stage.	
		(c) Consider use of a flow diagram.	
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders.	***Consider giving information by genotype**.*
		(b) Indicate number of participants with missing data for each variable of interest.	
		(b) **Cohort study –** Summarize follow-up time, e.g. average and total amount.	
Outcome data	15 *	**Cohort study-**Report numbers of outcome events or summary measures over time. **Case-control study –** Report numbers in each exposure category, or summary measures of exposure. **Cross-sectional study –** Report numbers of outcome events or summary	in cohort studies, ***report outcomes (phenotypes) for each genotype category over time;*** in case-control studies, ***report numbers in each genotype***

	Item number		*Extension for genetic association studies*
		measures. .	***category;*** in cross-sectional studies, ***report outcomes (phenotypes) for each genotype category.***
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence intervals). Make clear which confounders were adjusted for and why they were included.	
		(b) Report category boundaries when continuous variables were categorized.	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.	
			(d) Report results of any adjustments for multiple tests.

	Item number		*Extension for genetic association studies*
Other analyses	17	*(a)* Report other analyses done – e.g., analyses of subgroups and interactions, and sensitivity analyses.	
			(b) If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken. ***(c) If detailed results are available elsewhere, state how they can be accessed.***
DISCUSSION			
Key results	18	Summarize key results with reference to study objectives.	
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other	

	Item number		*Extension for genetic association studies*
		relevant evidence.	
Generalizability	21	Discuss the generalizability (external validity) of the study results.	
OTHER INFORMATION			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.	

* Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Table 2. New and modified items in STREGA (compared with STROBE), organized by topic

Specific issue in genetic association studies	**Rationale for inclusion in STREGA**	**Item(s) in STREGA**	**Comment**
Selection of participants	Selection bias may occur if (i) genetic associations are investigated in one or more subsets of participants (sub-samples) from a particular study; or (ii) there is differential non-participation in groups being compared; or, (iii) there are differential genotyping success on call rates in groups being compared.	6(a): [Give the eligibility criteria, and the sources and methods of selection of participants] *Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant*. 13(a): *Include numbers in whom genotyping was attempted and numbers in whom genotyping was successful.*	Inclusion and exclusion criteria, sources and methods of selection of sub-samples should be specified, stating whether these were based on a priori or post hoc considerations.
Rationale for choice of genes and variants investigated	Without an explicit rationale, it is difficult to judge the potential for selective reporting of study results. There is strong empirical evidence from randomised controlled trials that reporting of trial outcomes is frequently incomplete and	3: *State if the study is the first report of a genetic association, a replicated effort, or both.* 7(b): *Clearly define genetic exposures (genetic variants) using a widely-used nomenclature system. Identify variables likely to be*	The scientific background and rationale for investigating the genes and variants should be reported. For genome-wide association studies, it is important to specify what initial testing platforms were used and how gene variants are selected for further testing in subsequent stages

	biased in favour of statistically significant findings (145-147). Some evidence is also available in pharmacogenetics (148).	*associated with population stratification (confounding by ethnic origin).*	
Population stratification (confounding by ethnic origin)	When sub-populations are present in the study that differ both in allele (or genotype) frequencies and disease risks, then confounding will occur if these sub-populations are unevenly distributed across exposure groups (or between cases and controls). *See also main text.*	12(h): *Describe any methods used to assess or address population stratification.*	In view of the debate about the potential implications of population stratification for the validity of genetic association studies, transparent reporting of the methods used, or stating that none was used, to address this potential problem is important for allowing the empirical evidence to accrue. Ethnicity information should be presented (see for example Winker (149), as should genetic markers or other variables likely to be associated with population stratification. Details of case-family control designs should be provided if they are used. As several methods of adjusting for population stratification have been proposed (100), explicit documentation of the methods is needed.
Genotyping errors (misclassification of exposure)	Non-differential genotyping errors will usually bias associations towards the null (81, 82). When there are systematic differences in genotyping according to outcome status (differential error), bias in any direction may occur. *See also main text.*	8(b): *Describe specific laboratory methods, including source and storage of DNA, genotyping methods and platforms, error rates and call rates. State where genotyping was done.* 13(a): *Include numbers in whom genotyping was*	Factors affecting the potential extent of misclassification (information bias) of genotype include the types and quality of samples, timing of collection, and the method used for genotyping (18, 77,150). For case-control studies, whether or not genotyping was done blind to case-control status should be reported, along with the reason for this decision (see text).

		attempted and numbers in whom genotyping was successful.	
Haplotype inference	In designs considered in this paper, haplotypes have to be inferred because of lack of available family information. There are diverse methods for inferring haplotypes. *See also main text.*	12(g): *Describe any methods used for inferring genotypes or haplotypes.*	When discrete "windows" are used to summarize haplotypes, variation in the definition of these may complicate comparisons across studies, as results may be sensitive to choice of windows. Related "imputation" strategies are also in use (101, 127, 151). It is important to give details on haplotype inference and, when possible, uncertainty. Additional considerations for reporting include: the strategy for dealing with rare haplotypes, window size and construction (if used); choice of software.
Hardy-Weinberg equilibrium (HWE)	Departure from HWE may indicate errors or peculiarities in the data (118). Empirical assessments have found that 20-69% of genetic associations were reported with some indication about conformity with HWE, and that among some of these, there were limitations or errors in the assessment of HWE (118). *See also main text.*	12(f): *State whether Hardy-Weinberg equilibrium was considered and, if so, how.*	Any statistical tests or measures should be described, as should any procedure to allow for deviations from HWE in evaluating genetic associations (119).
Volume of data	The key problem is of possible false-positive findings and selective reporting of these. Type I	12(i): *Describe any methods used to address multiple tests or to control risk of false positive findings.*	GWA studies collect information on a very large number of genetic variants concomitantly. Initiatives to make the entire database transparent and available online may supply a definitive solution to the problem of selective reporting (7).

	errors are particularly relevant to the conduct of GWA studies. A large search among hundreds of thousands of genetic variants can be expected by chance alone to find thousands of false positive signals (risk ratios significantly different from 1.0).	16(d): *Report results of any adjustments for multiple tests.* 17(b): *If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken* 17(c): *If detailed results are available elsewhere, state how they can be accessed.*	The volume of data analyzed should also be considered in the interpretation of findings. Examples of methods of summarizing results include giving distribution of frequentist p values, distribution of effect sizes and specifying false discovery rates.
Reporting of data	The synthesis of findings across studies depends on the availability of sufficiently detailed data.	14(a): [Give characteristics of study participants] *Consider giving information by genotype.* 15: [For outcomes] *Provide outcomes (phenotypes) for each genotype category [over time*], or *Provide numbers in each genotype category.*	
Data analysis	Analysis methods should be transparent and replicable, and genetic association studies are often performed using specialized software.	12(a): [Describe all statistical methods…] *State software version and options (or settings) selected.*	

Table 3. Issues considered for potential inclusion in STREGA, but considered to be covered by STROBE

Issue	Where this is covered in STROBE	Comment
Quantitative outcomes Lack of information on the nature and circumstances of measurements may compromise comparison between studies, and meta-analysis.	Item 8	Although many genetic studies examine quantitative traits, it was concluded that this issue applies to observational studies in general
Exaggerated cases or controls (spectrum of disease bias) In some studies of continuous traits, cases have been over-selected from multiplex families with strong heritability and compared with controls from families without family history of the trait, or only very severely affected cases have been eligible and compared with controls selected from the opposite end of the distribution of the trait. If disease severity or heritability correlates with the strength of the association, then these studies may obtain different estimates of effects compared with studies that include a broader group unselected for disease severity or heritability.	Item 6	Measures of genetic effects tend to be higher in multiplex families than in the general population (152), but we lack empirical evidence on the extent of spectrum of disease bias in the genetic epidemiology of complex diseases.
Relatedness of study subjects Selection bias may occur when case-control genetic association studies include members of families previously ascertained, e.g., for genome linkage scans (153, 154). For example, a number of families initially collected for genome linkage scan studies	Items 6, 9, 12	Although more likely to be encountered in genetic context, it was concluded that this issue applies to observational studies in general and did not need additional coverage in STREGA. Cryptic relatedness of cases overlaps conceptually with

are now being used as platforms for GWA studies (155). Relatedness of study participants may be problematic in samples ascertained from isolated populations.		population stratification (100) with (unobserved) family being the potential confounder.
Joint effects (including gene-gene and gene-environment interaction) Possible selective reporting of results, risk of false negative and false positive tests of interaction.	Items 12(b), 17(a)	It was concluded that this issue applies to observational studies in general and did not need additional coverage in STREGA. The key issue is transparency of reporting a priori hypotheses, and then how many other tests for potential interactions were made.
Inference of causation in single studies Possible over-interpretation of results.	Item 20	It was concluded that this issue applies to observational studies in general and did not need additional coverage in STREGA. The general recommendation is to give a balanced overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. Claims for association should be tempered allowing for the extent of replication and protection from bias (39, 156). There has been considerable concern about non-replication of gene-disease association studies (12, 13, 157-163). However, this is a problem for observational studies in general (164, 165).

Appendix A: Participants in Workshop

Nick Birkett Associate Professor of Cancer and Epidemiology Department of Epidemiology and Community Medicine University of Ottawa, Faculty of Medicine 451 Smyth Road Ottawa, Ontario, Canada K1H 8M5 Phone: (613) 562 5800 ext 8289 E-mail: nbirkett@uOttawa.ca	**Michele Cargill** Associate Director, Affymetrix 3450 Central Expressway Santa Clara, California 95051 USA Phone: 408-731-5468 E-mail: Michele_Cargill@affymetrix.com
Barbara Cohen PLoS Medicine Public Library of Science 185 Berry Street, Suite 3100 San Francisco, CA 94107 USA Phone: (415) 624-1200 E-mail: bcohen@plos.org	**George Davey-Smith** International Journal of Epidemiology Professor of Clinical Epidemiology, Department of Social Medicine University of Bristol Canynge Hall, Whiteladies Road Bristol BS8 2PR England, U.K. Phone: +44 (0)117 928 7329 E-Mail: ije-editorial@bristol.ac.uk
Aaron Del Duca DNA Genotek 29 Camelot Dr. Unit 200 Ottawa, Ontario, Canada K2G 5W6 Phone: (613) 723 5757 ext 270	**Isabel Fortier** McGill University and Genome Innovation Center 740 Dr. Penfield Avenue Montréal, Québec, Canada, H3A 1A4 Tel: (514) 398 3311 x 00555 E-Mail: isabel.fortier@mail.mcgill.ca
Matthew Freedman* Postdoctoral Fellow, Center for Genome Research, MIT/Whitehead Genome Center/ Instructor, Department of Medicine, Harvard Medical School Massachusetts General Hospital Wellman Building Boston, MA, 02114 Mailstop: 8th floor E-mail: freedman@broad.mit.edu	**France Gagnon*** Assistant Professor of Genetic Epidemiology Department of Epidemiology and Community Medicine University of Ottawa, Faculty of Medicine 451 Smyth Road Ottawa, Ontario, Canada K1H 8M5 Phone: (613) 562 5800 ext 8262 E-mail: fgagnon@uottawa.ca
Jean Golding Emeritus Professor of Pediatric and Perinatal Epidemiology/ Founder and former Scientific and Executive Director, ALSPAC Community Based Medicine 24 Tyndall Avenue, Clifton Bristol BS8 1TQ England, UK Phone : +44 (0)117 928 5099 E-mail: Jean.Golding@bris.ac.uk	**Jeremy Grimshaw*** Director of Clinical Epidemiology Unit, Ontario Health Research Institute/ Professor, University of Ottawa 725 Parkdale Avenue Ottawa, Ontario, Canada K1Y 4E9 Phone: (613) 761 5231 E-mail: jgrimshaw@ohri.ca

Marta Gwinn* Medical Officer, Office of Genomics and Disease Prevention, Centers for Disease Control and Prevention (CDC) KOGR Building Atlanta, Georgia 30341 USA Phone: (770) 488 8433 E-mail: Mgwinn@cdc.gov	**Julian Higgins*** Research Associate, Medical Research Council (MRC) Biostatistics Unit MRC Biostatistics Unit Institute of Public Health Forvie Site Robinson Way Cambridge CB2 2SR England, U.K. Phone: +44 (0) 122 333 0376 E-mail: julian.higgins@mrc.bsu.cam.ac.uk
Claire Infante-Rivard Professor (James McGill Professor) McGill University Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine 1130 Pine Avenue West Montréal, Québec, Canada H3A 1A3 Phone : (514) 398 4231 E-mail: claire.infante-rivard@mcgill.ca	**John Ioannidis*** Associate Professor and Chairman, Department of Hygiene and Epidemiology, University of Ioannina Medical School P.O. Box 1186 45110 Ioannina, Greece Phone: +30 2651 097 1989 E-mail: jioannid@cc.uoi.gr
Richard King Institute for Human Genetics, University of Minnesota Department of Minnesota MMC 485 Mayo 420 Delaware St SE Minneapolis, MN 55455 USA Phone: (612) 624 6657 E-mail: kingx002@umn.edu	**Julian Little*** Professor of Molecular and Human Genome Epidemiology, Cancer Genetic Epidemiology, Genomics Health, Services Research and Genetic Epidemiology Department of Epidemiology and Community Medicine University of Ottawa, Faculty of Medicine 451 Smyth Road Ottawa, Ontario, Canada K1H 8M5 Phone: (613) 562 5800 ext 8689 E-mail: jlittle@uottawa.ca
John McLaughlin Senior Investigator and Head of Epidemiology and Biostatistics Samuel Lunenfeld Research Institute 60 Murray Street room 5-211 Toronto, Ontario, Canada M5G 1X5 Phone: (416) 586 4800 ext 8304 E-mail: jmclaugh@mshri.on.ca	**David Moher*** Director of Clinical Research at Children's Hospital of Eastern Ontario (CHEO) and Chalmers Research Group (CRG) 401 Smyth Road Ottawa, Ontario, Canada K1H 8L1 Phone: (613) 737 7600 ext 3956 E-mail: dmoher@uottawa.ca
Andrew Paterson Scientist, Program in Genetics & Genomic Biology The Hospital for Sick Children 555 University Avenue Toronto, Ontario, Canada M5G 1X8 Phone: (416) 813 6994 E-mail: Andrew.paterson@utoronto.ca	**Paul Scheet*** Postdoctoral Researcher, Department of Statistics The University of Washington Department of Statistics Box 354322 Seattle, Washington 98195-4322 USA Phone: (206) 543 7237 E-mail: paul@stat.washington.edu

Alexandre Stewart Principal Investigator, University of Ottawa Heart Institute 40 Ruskin Street, Ottawa, Ontario, Canada K1Y 4W7 Phone: (613) 761 5189 E-mail: astewart@ottawaheart.ca	**Cornelia van Duijn** Professor of Genetic Epidemiology Department of Epidemiology and Biostatistics Erasmus MC, Rotterdam P.O. Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31 10 408 7488 E-mail: c.vanduijn@erasmusmc.nl
Kyle Vogan Associate Editor, Nature Genetics Nature Genetics 75 Varick Street, 9th Floor New York NY 10013-1917 USA Phone: (212) 726 9314 E-mail: k.vogan@natureny.com	**Shalom Wacholder*** Division of Cancer Epidemiology and Genetics National Cancer Institute Executive Plaza South, Room 8046 Bethesda, MD 20892 USA Phone (301) 496 3358 E-mail: wacholds@exchange.nih.gov
George Wells Professor, The University of Ottawa/ Director, Cardiovascular Research Methods Centre Department of Epidemiology and Community Medicine/ University of Ottawa Heart Institute H1-1, 40 Ruskin Street Ottawa, Ontario K1Y 4W7 Phone: (613) 798 5555 ext 18640 E-mail: gawells@ottawaheart.ca	**Allen Wilcox** Senior Investigator, Division of Intramural Research Environmental Diseases & Medicine Program Epidemiology Branch NIEHS MD A3-05 P.O. BOX 12233 Research Triangle Park, NC 27709 USA Phone: (919) 541-4660 E-mail: wilcox@niehs.nih.gov
Robin Williamson Deputy Editor, The American Journal of Human Genetics Brigham and Women's Hospital New Research Building Room 160 77 Avenue Louis Pasteur Boston, MA 02115 USA Phone: (617) 525 4770 E-mail: rwilliamson@ajhg.net	**Maja Zecevic** Senior North American Editor The Lancet Publishing Group 360 Park Avenue South New York, NY 10010-1710 USA E-mail: m.zecevic@lancet.com
Guangyong Zou* Scientist CRTG, Clinical Trials Robarts Research Institute 1, 1st floor PO BOX 5015 100 Perth Drive London, Ontario, Canada N6A 5K8 Phone: (519) 663 3400 ext 34092 E-mail: gzou@robarts.ca	

* indicates speaker

Rapporteurs

Kimberly Hutchings University of Ottawa Department of Epidemiology and Community Medicine 451 Smyth Rd., Room 3204 Ottawa, Ontario, Canada K1H 8M5 E-mail: hutchingskimberley@rogers.com	**Val Tait** University of Ottawa Department of Epidemiology and Community Medicine 451 Smyth Rd., Room 3204 Ottawa, Ontario, Canada K1H 8M5 E-mail: vtait077@uottawa.ca
Candice Yang Johnson University of Ottawa Department of Epidemiology and Community Medicine 451 Smyth Rd., Room 3204 Ottawa, Ontario, Canada K1H 8M5 E-mail: candice.yang.johnson@gmail.com	**Miriam Wiens** University of Ottawa Department of Epidemiology and Community Medicine 451 Smyth Rd., Room 3204 Ottawa, Ontario, Canada K1H 8M5 E-mail: mir_wiens@hotmail.com

STREGA Workshop Coordinator

Silvia Visentin
University of Ottawa
Department of Epidemiology and Community Medicine
451 Smyth Rd., Room 3204
Ottawa, Ontario, Canada K1H 8M5
Phone: (613) 562 5800 ext 8719
E-mail: svisenti@uottawa.ca

Appendix B: Chinese journals that have published genetic association studies, indexed in PubMed: Rank and number of genetic association studies published, October 2000-December 2005

Journal	Rank (appproximate§ among 1609)	Number of publications
Zhonghua Yi Xue Yi Chuan Xue Za Zhi	5	223
Zhonghua Yi Xue Za Zhi	33	96
Zhonghua Liu Xing Bing Xue Za Zhi	82	55
Zhonghua Nei Ke Za Zhi	148	33
Di Yi Jun Yi Da Xue Xue Bao	170	28
Zhonghua Fu Chan Ke Za Zhi	175	26
Zhonghua Yu Fang Yi Xue Za Zhi	186	24
Yi Chuan Xue Bao	187	24
Zhongguo Yi Xue Ke Xue Yuan Xue Bao	195	22
Zhonghua Xue Ye Xue Za Zhi	220	19
Zhonghua Jie He He Hu Xi Za Zhi	221	19
Zhongguo Shi Yan Xue Ye Xue Za Zhi	234	18
Wei Sheng Yan Jiu	235	18
Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi	242	17
Zhonghua Zhong Liu Za Zhi	252	16
Zhonghua Xin Xue Guan Bing Za Zhi	264	15
Zhonghua Er Ke Za Zhi	278	14
J Huazhong Univ Sci Technolog Med Sci	285	14
Ai Zheng	293	14
Zhonghua Gan Zang Bing Za Zhi	314	12
Yonsei Med J	315	12
Zhonghua Wai Ke Za Zhi	469	7
Zhonghua Kou Qiang Yi Xue Za Zhi	470	7
Zhonghua Er Bi Yan Hou Ke Za Zhi	509	6
Zhonghua Nan Ke Xue	558	5
Zhongguo Wei Zhong Bing Ji Jiu Yi Xue	559	5
Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)	564	5
Fa Yi Xue Za Zhi	604	5
Zhonghua Zheng Xing Wai Ke Za Zhi	626	4
Zhong Nan Da Xue Xue Bao Yi Xue Ban	627	4
Yao Xue Xue Bao	628	4
Shanghai Kou Qiang Yi Xue	637	4
J Tongji Med Univ	673	4
Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi	729	3

Zhongguo Zhong Xi Yi Jie He Za Zhi	730	3
Zhejiang Da Xue Xue Bao Yi Xue Ban	731	3
Hua Xi Kou Qiang Yi Xue Za Zhi	814	3
Zhonghua Yan Ke Za Zhi	873	2
Zhonghua Bing Li Xue Za Zhi	874	2
Yan Ke Xue Bao	876	2
Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi	877	2
Zhongguo Ying Yong Sheng Li Xue Za Zhi	1091	1
Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi	1092	1
Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi	1093	1
Zhong Xi Yi Jie He Xue Bao	1094	1

§ Approximate as when there is a tie, the journals are listed in reverse alphabetical order.

The source of this information was the HuGE Published Literature database (3, 4), which is built from references extracted from PubMed. This extraction process was started in October 2000 and involves review of abstracts for relevance to human genome epidemiology (specifically studies with information on one or more of genotype prevalence, gene- disease associations, gene-environment or gene-gene interactions, or evaluations of genetic tests). If relevance is unclear from the abstract, the full paper is checked.

Appendix C: STREGA Workshop Agenda

Day 1 – Thursday, June 15, 2006

Time	Agenda – Day 1 Thursday, June 15, 2006	Speakers
7:00 – 7:50 am	Breakfast (Beau Rivage B Foyer) Registration (Frontenac Meeting Room)	
8:00 – 8:15 am	Introduction to workshop – objectives procedures and anticipated outputs	**Julian Little**
8:15 – 8:30 am	Roadmap for efficient and reliable human genome epidemiology *This workshop is intended to precede one on grading of evidence from genetic association studies, and application of this to field synopses*	**John Ioannidis**
8:30 – 9:00 am	Overview of existing guidance on reporting of genetic association studies – discussion of epidemiological, genetic and journal editor perspectives. *Extent to which evidence based; strengths and limitations; impact*	**Julian Little**
9:00 – 9:30 am	Empirical evidence on reporting of genetic association studies: preliminary results of analysis of 5% sample of HuGE published literature database	**Muin Khoury** **John Ioannidis**
9:30 – 10:00 am	The STROBE statement *Explanation of process of development of STROBE, perceived strengths and gaps*	**Julian Little on behalf of Eric von Elm**
10:00 – 10:30 am	Nutrition Break (Frontenac Foyer)	
10:30 – 11:30 am	**Cross-cutting issues especially pertinent to genetic association studies –part 1**	
10:30-10:50 am	Continuous traits (including extreme vs. unselected cases)	**France Gagnon**
10:50-11:10 am	Reporting of selection and participation of study subjects *This is a general issue in observational studies. Impact on genetic association studies not well understood*	**Julian Higgins**
11:10-11:30 am	Reporting of rationale for choice of genes and variants investigated	**Julian Little**
11:30 – 11:40 am	Comfort Break	
11:40 am -12:00	Reporting of genotyping methods: (i) genotyping error	**France Gagnon**
12:00-12:20 pm	(ii) haplotypes	**Paul Scheet**

	reporting of these in way that information can be combined across studies may not be easy to sort out	
12:20-12:30 pm	General discussion	
12:30 – 1:30 pm	Lunch Break (Beau Rivage B)	
1:30 – 2:45 pm	**Cross-cutting issues especially pertinent to genetic association studies –part 2: population stratification & joint effects**	
1:30-2:00 pm	Reporting potential for population stratification: genetic and epidemiological perspectives	**Matthew Freedman Sholom Wacholder**
2:00-2:45 pm	Reporting of gene-environment and gene-gene interaction	**Marta Gwinn on behalf of Muin Khoury**
2:45-3:00 pm	Assignment for breakout groups	
3:00 – 3:15 pm	Nutrition Break (Frontenac Foyer)	
3:15 – 4:30	Study design-specific breakout sessions to develop checklists for reporting genetic association studies Three groups: 1. Cohort studies (Location #1) 2. Case- control studies (Location #2) 3. Cross-sectional studies (Location #3) *We propose to involve graduate students in the meeting; one contribution would be to act as rapporteurs to these groups* *Chairs of groups identified in advance of meeting*	
4:30 – 4:40 pm	Comfort Break	
4:40 – 5:40 pm	Breakout groups continue work; prioritise issues for next day, and summary of issues to raise at plenary that opens day 2	
7:00 pm	**BBQ Dinner – Hotel (Champion Room)**	

Day 2 – Friday, June 16, 2006

Time	Agenda – Day 2 Friday, June 16, 2006	Speakers
7:00 – 7:50 am	Breakfast (Beau Rivage B Foyer)	
8:00 – 8:15 am	Plenary session; procedures for the day; brief issues from breakout sessions	
8:15 – 9:30 am	**Cross-cutting issues especially pertinent to genetic association studies –part 3: HWE, statistical analysis and interpretation**	
8:15-8:40 am	Hardy-Weinberg equilibrium	**Guang Yong Zou**
8:40-9:05 am	Multiple testing and pre-study odds of true finding; selective reporting *especially pertinent to statistical analysis of genetic association studies*	**Sholom Wacholder**
9:05-9:45 am	Inference of causation in single studies *Exaggerated claims have been made for single studies on the basis of limited mechanistic evidence; is there scope for an evidence-based plea for moderation?!*	**John Ioannidis**
9:45 – 10:15 am	Nutrition Break (Frontenac Foyer)	
10:15 – 12:15 pm	Study design-specific breakout sessions to develop checklists for reporting genetic association studies; preparation of conclusions to take to plenary after lunch	
12:15 – 1:15 pm	Lunch (Frontenac Foyer)	
1:15 – 2:15 pm	Study design-specific breakout sessions to prepare conclusions to take to plenary	
2:15 – 3:00 pm	Presentation of conclusions from break-out groups, and discussion	
3:00 – 3:30 pm	Nutrition Break (Frontenac Foyer)	
3:30 pm – 5:00 pm	**Conclusions**	
3.30-3.50 pm	Gaps in the evidence	
3.50-4.10 pm	Dissemination	**Jeremy Grimshaw**
4.10-4.40 pm	Plans for evaluating impact of STREGA	**David Moher**
4.40-5.00 pm	Next steps	

References

1. Khoury MJ, Little J, Burke W. Human Genome Epidemiology: Scope and Strategies. In: Khoury MJ, Little J, Burke W, eds. Human Genome Epidemiology New York: Oxford University Press, 2004:3-16.

2. Genomics, Health and Society Working Group. ***Genomics, Health and Society. Emerging Issues for Public Policy***. Ottawa: Government of Canada Policy Research Initiative, 2004.

3. Lin BK, Clyne M, Walsh M, Gomez O, Yu W, Gwinn M, Khoury MJ. Tracking the Epidemiology of Human Genes in the Literature: The HuGE Published Literature Database. Am.J.Epidemiol. 2006; 164:1-4.

4. Yu Y, Yesupriya A, Clyne M, Wulf A, Gwinn M, Khoury MJ. HuGE Literature Finder. HuGE Navigator. Available at http://www.hugenavigator.net/HuGENavigator/startPagePubLit.do/. Accessed 30 January 2008.

5. Lawrence RW, Evans DM, Cardon LR. Prospects and pitfalls in whole genome association studies. Philos.Trans.R.Soc.Lond.B.Biol.Sci. 2005; 360:1589-1595.

6. Thomas DC. Are we ready for genome-wide association studies? Cancer Epidemiol.Biomarkers Prev. 2006; 15:595-598.

7. Khoury MJ, Little J, Gwinn M, Ioannidis JP. On the synthesis and interpretation of consistent but weak gene-disease associations in the era of genome-wide association studies. Int.J.Epidemiol. 2007; 36:439-445.

8. Little J, Khoury MJ, Bradley L, Clyne M, Gwinn M, Lin B, Lindegren ML, Yoon P. The human genome project is complete. How do we develop a handle for the pump? Am J Epidemiol 2003; 157:667-673.

9. Ioannidis JP, Bernstein J, Boffetta P, Danesh J, Dolan S, Hartge P, Hunter D, Inskip P, Jarvelin MR, Little J, Maraganore DM, Bishop JA, O'Brien TR, Petersen G, Riboli E, Seminara D, Taioli E, Uitterlinden AG, Vineis P, Winn DM, Salanti G, Higgins JP, Khoury MJ. A network of investigator networks in human genome epidemiology. Am.J.Epidemiol. 2005; 162:302-304.

10. Ioannidis JP, Gwinn M, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam A, Danesh J, Smith GD, Dolan S, Duncan R, Gruis NA, Hartge P, Hashibe M, Hunter DJ, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, O'Brien TR, Petersen G, Riboli E, Salanti G, Seminara D, Smeeth L, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Khoury MJ, Human Genome Epidemiology Network and the Network of Investigator Networks. A road map for efficient and reliable human genome epidemiology. Nat.Genet. 2006; 38:3-5.

11. von Elm E, Egger M. The scandal of poor epidemiological research. BMJ 2004; 329:868-869.

12. Nature Genetics. Freely associating (editorial). Nat Genet 1999; 22:1-2.

13. Cardon L, Bell J. Association study designs for complex diseases. Nat Rev Genet 2001; 2:91-99.

14. Weiss S. Association studies in asthma genetics. American Journal of Respiratory and Critical Care Medicine 2001; 164:2014-2015.

15. Weiss ST, Silverman EK, Palmer LJ. Case-control association studies in pharmacogenetics. Pharmacogenomics J. 2001; 1:157-158.

16. Cooper DN, Nussbaum RL, Krawczak M. Proposed guidelines for papers describing DNA polymorphism-disease associations. Hum Genet 2002; 110:208.

17. Hegele R. SNP judgements and freedom of association. Arterioscler. Thromb. Vasc. Biol. 2002; 22:1058-1061.

18. Little J, Bradley L, Bray MS, Clyne M, Dorman J, Ellsworth DL, Hanson J, Khoury M, Lau J, O'Brien TR, Rothman N, Stroup D, Taioli E, Thomas D, Vainio H, Wacholder S, Weinberg C. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol 2002; 156:300-310.

19. Romero R, Kuivaniemi H, Tromp G, Olson JM. The design, execution, and interpretation of genetic association studies to decipher complex diseases. Am J Obstet Gynecol 2002; 187:1299-1312.

20. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet 2003; 361:865-872.

21. van Duijn CM, Porta M. Good prospects for genetic and molecular epidemiologic studies in the European Journal of Epidemiology. Eur.J.Epidemiol. 2003; 18:285-286.

22. Crossman D, Watkins H. Jesting Pilate, genetic case-control association studies, and Heart. Heart 2004; 90:831-832.

23. Huizinga TW, Pisetsky DS, Kimberly RP. Associations, populations, and the truth: recommendations for genetic association studies in Arthritis & Rheumatism. Arthritis & Rheumatism 2004; 50:2066-2071.

24. Little J. Reporting and Review of Human Genome Epidemiology Studies. In: Khoury MJ, Little J, Burke W, eds. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease New York: Oxford University Press, 2004:168-192.

25. Rebbeck TR, Martinez ME, Sellers TA, Shields PG, Wild CP, Potter JD. Genetic variation and cancer: improving the environment for publication of association studies. Cancer Epidemiology, Biomarkers & Prevention 2004; 13:1985-1986.

26. Tan N, Mulley J, Berkovic S. Association studies in epilepsy: "the truth is out there". Epilepsia 2004; 45:1429-1442.

27. Anonymous Framework for a fully powered risk engine. Nat.Genet. 2005; 37:1153.

28. Ehm MG, Nelson MR, Spurr NK. Guidelines for conducting and reporting whole genome/large-scale association studies. Hum.Mol.Genet. 2005; 14:2485-2488.

29. Freimer NB, Sabatti C. Guidelines for association studies in Human Molecular Genetics. Hum.Mol.Genet. 2005; 14:2481-2483.

30. Hattersley AT, McCarthy MI. What makes a good genetic association study? Lancet 2005; 366:1315-1323.

31. Manly K. Reliability of statistical associations between genes and disease. Immunogenetics 2005; 57:549-558.

32. Shen H, Liu Y, Liu P, recker R, Deng H. Nonreplication in genetic studies of complex diseases - lessons learned from studies of osteoporosis and tentative remedies. Journal of Bone and Mineral Research 2005; 20:365-376.

33. Vitali S, Randolph A. Assessing the quality of case-control association studies on the genetic basis of sepsis. Pediatr Crit Care Med 2005; 6:S74-S77.

34. Wedzicha JA, Hall IP. Publising genetic association studies in *Thorax*. Thorax 2005; 60:357.

35. Hall IP, Blakey JD. Genetic association studies in Thorax. Thorax 2005; 60:357-359.

36. DeLisi LE, Faraone SV. When is a "positive" association truly a "positive" in psychiatric genetics? A commentary based on issues debated at the World Congress of Psychiatric Genetics, Boston, October 12-18, 2005.
Am.J.Med.Genet.B.Neuropsychiatr.Genet. 2006; 141:319-322.

37. Saito YA, Talley NJ, de Andrade M, Petersen GM. Case-control genetic association studies in gastrointestinal disease: review and recommendations. Am.J.Gastroenterol. 2006; 101:1379-1389.

38. Uhlig K, Menon V, Schmid CH. Recommendations for reporting of clinical research studies. Am.J.Kidney Dis. 2007; 49:3-7.

39. NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P,

Fraumeni JF,Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype-phenotype associations. Nature 2007; 447:655-660.

40. Begg CB. Reflections on publication criteria for genetic association studies. Cancer Epidemiol.Biomarkers Prev. 2005; 14:1364-1365.

41. Byrnes G, Gurrin L, Dowty J, Hopper JL. Publication policy or publication bias? Cancer Epidemiol.Biomarkers Prev. 2005; 14:1363.

42. Pharoah PD, Dunning AM, Ponder BA, Easton DF. The reliable identification of disease-gene associations. Cancer Epidemiol.Biomarkers Prev. 2005; 14:1362.

43. Wacholder S. Publication environment and broad investigation of the genome. Cancer Epidemiol.Biomarkers Prev. 2005; 14:1361.

44. Whittemore AS. Genetic association studies: time for a new paradigm? Cancer Epidemiol.Biomarkers Prev. 2005; 14:1359-1360.

45. Bogardus ST,Jr., Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research. The need for methodological standards. JAMA 1999; 281:1919-1926.

46. Peters DL, Barber RC, Flood EM, Garner HR, O'Keefe GE. Methodologic quality and genotyping reproducibility in studies of tumor necrosis factor -308 G-->A single nucleotide polymorphism and bacterial sepsis: implications for studies of complex traits. Crit.Care Med. 2003; 31:1691-1696.

47. Clark MF, Baudouin SV. A systematic review of the quality of genetic association studies in human sepsis. Intensive Care Med. 2006; 32:1706-1712.

48. Lee W, Bindman J, Ford T, Glozier N, Moran P, Stewart R, Hotopf M. Bias in psychiatric case-control studies: literature survey. Br.J.Psychiatry 2007; 190:204-209.

49. Yesupriya AJ, Evangelou E, Kavvoura FK, Patsopoulos NA, Clyne M, Walsh M, Lin BK, Yu W, Gwinn M, Ioannidis JPA, Khoury MJ. Reporting of human genome epidemiology (HuGE) association studies: an empirical assessment. BMC Medical Genetics 2007 (under review)

50. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. JAMA 1995; 274:645-651.

51. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-kremer S, Stewart J, Taylor R, Vilo J, Vingron M. Minimum information about a microarray

experiment (MIAME) - toward standards for microarray data. Nat.Genet. 2001; 29:356-371.

52. Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE, McCormack VA. Issues in the reporting of epidemiological studies: a survey of recent practice. BMJ 2004; 329:883.

53. Altman D, Moher D. Developing guidelines for reporting healthcare research: scientific rationale and procedures. Med Clin (Barc) 2005; 125:8-13.

54. Gluud LL. Bias in clinical intervention research. Am.J.Epidemiol. 2006; 163:493-501.

55. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gotzsche PC, Lang T, CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann.Intern.Med. 2001; 134:663-694.

56. Moher D, Schultz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001; 285:1987-1991.

57. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, Gaboury I. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med.J.Aust. 2006; 185:263-267.

58. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann.Intern.Med. 2007; 147:573-577.

59. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. B.M.J 2007; 335:806-808.

60. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull.World Health Organ. 2007; 85:867-872.

61. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology 2007; 18:800-804.

62. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in

Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007; 370:1453-1457.

63. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. Prev.Med. 2007; 45:247-251.

64. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. PLoS Med. 2007; 4:e296.

65. Ebrahim S, Clarke M. STROBE: new standards for reporting observational epidemiology, a chance to improve. Int.J.Epidemiol. 2007; 36:946-948.

66. Egger M, Altman DG, Vandenbroucke JP, of the STROBE group. Commentary: Strengthening the reporting of observational epidemiology the STROBE statement. Int.J.Epidemiol. 2007; 36:948-950.

67. Editors. Probing STROBE. Epidemiology 2007; 18:789-790.

68. Vandenbroucke JP. The making of STROBE. Epidemiology 2007; 18:797-799.

69. Rothwell PM, Bhatia M. Reporting of observational studies. BMJ 2007; 335:783-784.

70. Little J, Higgins JPT (editors). The HuGENet™ HuGE Review Handbook, version 1.0. Available at http://www.hugenet.ca. Accessed 30 January 2008.

71. Higgins JP, Little J, Ioannidis JP, Bray MS, Manolio TA, Smeeth L, Sterne JA, Anagnostelis B, Butterworth AS, Danesh J, Dezateux C, Gallacher JE, Gwinn M, Lewis SJ, Minelli C, Pharoah PD, Salanti G, Sanderson S, Smith LA, Taioli E, Thompson JR, Thompson SG, Walker N, Zimmern RL, Khoury MJ. Turning the Pump Handle: Evolving Methods for Integrating the Evidence on Gene-Disease Association. Am.J.Epidemiol. 2007; 166:863-866.

72. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Ann.Intern.Med. 2007; 147:W163-94.

73. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology 2007; 18:805-835.

74. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, for the STROBE Initiative. Strengthening the

Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Med. 2007; 4:e297.

75. Rothman KJ, Poole C. Some guidelines on guidelines. They should come with expiration dates. Epidemiology 2007; 18:794-796.

76. Kuller LH, Goldstein BD. Suggestions for STROBE recommendations. Epidemiology 2007; 18:792-793.

77. Pompanon F, Bonin A, Bellemain E, Taberlet P. Genotyping errors: causes, consequences and solutions. Nat.Rev.Genet. 2005; 6:847-859.

78. Akey JM, Zhang K, Xiong M, Doris P, Jin L. The effect that genotyping errors have on the robustness of common linkage-disequilibrium measures. Am.J.Hum.Genet. 2001; 68:1447-1456.

79. Dequeker E, Ramsden S, Grody WW, Stenzel TT, Barton DE. Quality control in molecular genetic testing. Nature Reviews Genetics 2001; 2:717-723.

80. Mitchell AA, Cutler DJ, Chakravarti A. Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. Am.J.Hum.Genet. 2003; 72:598-610.

81. Rothman N, Stewart WF, Caporaso NE, Hayes RB. Misclassification of genetic susceptibility biomarkers: implications for case-control studies and cross-population comparisons. Cancer Epidemiol Biomarkers Prev 1993; 2:299-303.

82. Garcia-Closas M, Wacholder S, Caporaso N, Rothman N. Inference Issues in Cohort and Case-Control Studies of Genetic Effects and Gene-Environment Interactions. In: Khoury MJ, Little J, Burke W, eds. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease New York: Oxford University Press, 2004:127-144.

83. Wong MY, Day NE, Luan JA, Wareham NJ. Estimation of magnitude in gene-environment interactions in the presence of measurement error. Stat.Med. 2004; 23:987-998.

84. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD, Todd JA. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat.Genet. 2005; 37:1243-1246.

85. Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. Am J Human Genet 1988; 43:520-526.

86. Gelernter J, Goldman D, Risch N. The A1 allele at the D2 dopamine receptor gene and alcoholism: a reappraisal. JAMA 1993; 269:1673-1677.

87. Kittles RA, Chen W, Panguluri RK, Ahaghotu C, Jackson A, Adebamowo CA, Griffin R, Williams T, Ukoli F, Adams-Campbell L, Kwagyan J, Isaacs W, Freeman V, Dunston GM. CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? Hum Genet 2002; 110:553-560.

88. Thomas DC, Witte JS. Point: population stratification: a problem for case control studies of candidate-gene associations? Cancer Epidemiol Biomarkers Prev 2002; 11:505-512.

89. Wacholder S, Chatterjee N, Hartge P. Joint effects of genes and environment distorted by selection biases: implications for hospital-based case-control studies. Cancer Epidemiol Biomarkers Prev 2002; 11:885-889.

90. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet 2003; 361:598-604.

91. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2000; 92:1151-1158.

92. Ardlie KG, Lunetta KL, Seielstad M. Testing for population subdivision and association in four case-control studies. Am J Human Genet 2002; 71:304-311.

93. Edland SD, Slager S, Farrer M. Genetic association studies in Alzheimer's disease research: challenges and opportunities. Stat.Med. 2004; 23:169-178.

94. Millikan RC. Re: population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2001; 93:156-157.

95. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. Genet Epidemiol 2004; 27:14-20.

96. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. Nat.Genet. 2004; 36:1312-1318.

97. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. Nat Genet 2004; 36:512-517.

98. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D. Assessing the impact of population stratification on genetic association studies. Nat Genet 2004; 36:388-393.

99. Khlat M, Cazes MH, Genin E, Guiguet M. Robustness of case-control studies of genetic factors to population stratification: magnitude of bias and type I error. Cancer Epidemiol.Biomarkers Prev. 2004; 13:1660-1664.

100. Balding DJ. A tutorial on statistical methods for population association studies. Nat.Rev.Genet. 2006; 7:781-791.

101. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447:661-678.

102. Ioannidis JP. Non-Replication and Inconsistency in the Genome-Wide Association Setting. Hum.Hered. 2007; 64:203-213.

103. Zhao LP, Li SS, Khalid N. A method for the assessment of disease associations with single-nucleotide polymorphism haplotypes and environmental variables in case-control studies. Am J Hum Genet 2003; 72:1231-1250.

104. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z,

Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449:851-861.

105. Huang Q, Fu YX, Boerwinkle E. Comparison of strategies for selecting single nucleotide polymorphisms for case/control association studies. Hum Genet 2003; 113:253-257.

106. Kamatani N, Sekine A, Kitamoto T, Iida A, Saito S, Kogame A, Inoue E, Kawamoto M, Harigari M, Nakamura Y. Large-scale single-nucleotide polymorphism (SNP) and haplotype analyses, using dense SNP maps, of 199 drug-related genes in 752 subjects: the analysis of the association between uncommon SNPs within haplotype blocks and the haplotypes constructed with haplotype-tagging SNPs. Am J Hum Genet 2004; 75:190-203.

107. Zhang W, Collins A, Morton NE. Does haplotype diversity predict power for association mapping of disease susceptibility? Hum Genet 2004; 115:157-164.

108. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analysis using linkage disequilibrium. Am J Hum Genet 2004; 74:106-120.

109. van Hylckama Vlieg A, Sandkuijl LA, Rosendaal FR, Bertina RM, Vos HL. Candidate gene approach in association studies: would the factor V Leiden mutation have been found by this approach? Eur J Hum Genet 2004; 12:478-482.

110. Cardon LR, Abecasis GR. Using haplotype blocks to map human complex triat loci. Trends Genet 2003; 19:135-140.

111. Ke X, Hunt S, Tapper W, Lawrence R, Stavrides G, Ghori J, Whittaker P, Collins A, Morris AP, Bentley D, Cardon LR, Deloukas P. The impact of SNP density on fine-scale patterns of linkage disequilibrium. Hum Mol Genet 2004; 13:577-588.

112. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am.J.Hum.Genet. 2006; 78:629-644.

113. Servin B, Stephens M. Imputation-based analysis of association studies: candidate regions and quantitative traits. PLoS Genet. 2007; 3:e114.

114. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat.Genet. 2007; 39:906-913.

115. Hardy GH. Mendelian proportions in a mixed population. Science 1908; 28:49-50.

116. Xu J, Turner A, Little J, Bleecker ER, Meyers DA. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? Hum.Genet. 2002; 111:573-574.

117. Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, Riley J, Purvis I, Xu CF. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. European Journal of Human Genetics 2004; 12:395-399.

118. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. Eur.J.Hum.Genet. 2005; 13:840-848.

119. Zou GY, Donner A. The merits of testing Hardy-Weinberg equilibrium in the analysis of unmatched case-control data: a cautionary note. Ann.Hum.Genet. 2006; 70:923-933.

120. Shoemaker J, Painter I, Weir BS. A Bayesian characterization of Hardy-Weinberg disequilibrium. Genetics 1998; 149:2079-2088.

121. Ayres KL, Balding DJ. Measuring departures from Hardy-Weinberg: a Markov chain Monte Carlo method for estimating the inbreeding coefficient. Heredity 1998; 80 (Pt 6):769-777.

122. Trikalinos TA, Salanti G, Khoury MJ, Ioannidis JP. Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations. Am.J.Epidemiol. 2006; 163:300-309.

123. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D, Drummond H, Lees CW, Khawaja SA, Bagnall R, Burke DA, Todhunter CE, Ahmad T, Onnie CM, McArdle W, Strachan D, Bethel G, Bryan C, Lewis CM, Deloukas P, Forbes A, Sanderson J, Jewell DP, Satsangi J, Mansfield JC, the Wellcome Trust Case Control Consortium, Cardon L, Mathew CG. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat.Genet. 2007; 39:830-832.

124. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F, Lowe CE, Szeszko JS, Hafler JP, Zeitels L, Yang JH, Vella A, Nutland S, Stevens HE, Schuilenburg H, Coleman G, Maisuria M, Meadows W, Smink LJ, Healy B, Burren OS, Lam AA, Ovington NR, Allen J, Adlem E, Leung HT, Wallace C, Howson JM, Guja C, Ionescu-Tirgoviste C, Genetics of Type 1 Diabetes in Finland, Simmonds MJ, Heward JM, Gough SC, Dunger DB, the Wellcome Trust Case Control Consortium, Wicker LS, Clayton DG. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat.Genet. 2007; 39:857-864.

125. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M,

Hitman GA, Morris AD, Doney AS, Wellcome Trust Case Control Consortium (WTCCC), Burton PR, Clayton DG, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Davison D, Easton D, Evans D, Leung HT, Spencer CC, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop DT, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Mathew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop GM, Connell J, Dominiczak A, Braga Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hider SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Dunger DB, Widmer B, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AV, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Seal S, Stratton MR, Rahman N, Ban M, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghori MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widden C, Withers D, Cardin NJ, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Su Z, Teo YY, Vukcevic D, Bentley D, Compston A, Ouwehand NJ, Samani MR, Isaacs JD, Morgan AW, Wilson GD, Ardern-Jones A, Berg J, Brady A, Bradshaw N, Brewer C, Brice G, Bullman B, Campbell J, Castle B, Cetnarsryj R, Chapman C, Chu C, Coates N, Cole T, Davidson R, Donaldson A, Dorkins H, Douglas F, Eccles D, Eeles R, Elmslie F, Evans DG, Goff S, Goodman S, Goudie D, Gray J, Greenhalgh L, Gregory H, Hodgson SV, Homfray T, Houlston RS, Izatt L, Jackson L, Jeffers L, Johnson-Roffey V, Kavalier F, Kirk C, Lalloo F, Langman C, Locke I, Longmuir M, Mackay J, Magee A, Mansour S, Miedzybrodzka Z, Miller J, Morrison P, Murday V, Paterson J, Pichert G, Porteous M, Rahman N, Rogers M, Rowe S, Shanley S, Saggar A, Scott G, Side L, Snadden L, Steel M, Thomas M, Thomas S, McCarthy MI, Hattersley AT. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007; 316:1336-1341.

126. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J,

Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331-1336.

127. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007; 316:1341-1345.

128. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, Jonasdottir A, Jonasdottir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiasdottir S, Jonsdottir T, Palsson S, Einarsdottir H, Gunnarsdottir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdottir U, Kong A, Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007; 316:1491-1493.

129. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC. A common allele on chromosome 9 associated with coronary heart disease. Science 2007; 316:1488-1491.

130. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, SEARCH collaborators, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odefrey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo KY, Noh DY, Ahn SH, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low YL, Bogdanova N, Schurmann P, Dork T, Tollenaar RA, Jacobi CE, Devilee P, Klijn JG, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MW, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U, Ko YD, Spurdle AB, Beesley J, Chen X, kConFab, AOCS Management Group, Mannermaa A, Kosma VM, Kataja V, Hartikainen J, Day NE, Cox DR, Ponder BA. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007; 447:1087-1093.

131. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC,

Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF,Jr, Hoover RN, Thomas G, Chanock SJ. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat.Genet. 2007; 39:870-874.

132. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, Frigge ML, Geller F, Gudbjartsson D, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, Johannsson J, Sveinsson T, Myrdal G, Grimsson HN, Jonsson T, von Holst S, Werelius B, Margolin S, Lindblom A, Mayordomo JI, Haiman CA, Kiemeney LA, Johannsson OT, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat.Genet. 2007; 39:865-869.

133. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat.Genet. 2007; 39:977-983.

134. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, Waliszewska A, Neubauer J, Tandon A, Schirmer C, McDonald GJ, Greenway SC, Stram DO, Le Marchand L, Kolonel LN, Frasco M, Wong D, Pooler LC, Ardlie K, Oakley-Girvan I, Whittemore AS, Cooney KA, John EM, Ingles SA, Altshuler D, Henderson BE, Reich D. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat.Genet. 2007; 39:638-644.

135. Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang Z, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF,Jr, Hoover R, Hunter DJ, Chanock SJ, Thomas G. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat.Genet. 2007; 39:645-649.

136. Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM, Prendergast J, Olschwang S, Chiang T, Crowdy E, Ferretti V, Laflamme P, Sundararajan S, Roumy S, Olivier JF, Robidoux F, Sladek R, Montpetit A, Campbell P, Bezieau S, O'shea AM, Zogopoulos G, Cotterchio M, Newcomb P, McLaughlin J, Younghusband B, Green R, Green J, Porteous ME, Campbell H, Blanche H, Sahbatou M, Tubacher E, Bonaiti-Pellie C, Buecher B, Riboli E, Kury S, Chanock SJ, Potter J, Thomas G, Gallinger S, Hudson TJ, Dunlop MG. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat.Genet. 2007; 39:989-994.

137. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, Penegar S, Chandler I, Gorman M, Wood W, Barclay E, Lubbe S, Martin L, Sellick G, Jaeger E, Hubner R, Wild R, Rowan A, Fielding S, Howarth K, the CORGI Consortium, Silver A, Atkin W, Muir K, Logan R, Kerr D, Johnstone E, Sieber O, Gray R, Thomas H, Peto J, Cazier JB, Houlston R. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat.Genet. 2007; 39:984-988.

138. Haiman CA, Le Marchand L, Yamamoto J, Stram DO, Sheng X, Kolonel LN, Wu AH, Reich D, Henderson BE. A common genetic risk factor for colorectal and prostate cancer. Nature Genetics 2007; 39:954-956.

139. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. Science 2005; 308:385-389.

140. Haines JL, Hauser MA, Schmidt S, Scott WK, Olson LM, Gallins P, Spencer KL, Kwan SY, Noureddine M, Gilbert JR, Schnetz-Boutaud N, Agarwal A, Postel EA, Pericak-Vance MA. Complement factor H variant increases the risk of age-related macular degeneration. Science 2005; 308:419-421.

141. Edwards AO, Ritter R,3rd, Abel KJ, Manning A, Panhuysen C, Farrer LA. Complement factor H polymorphism and age-related macular degeneration. Science 2005; 308:421-424.

142. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard EJ, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhart AH, Rotter JI, Duerr RH, Cho JH, Daly MJ, Brant SR. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat.Genet. 2007; 39:596-604.

143. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, Vermeire S, Dewit O, de Vos M, Dixon A, Demarche B, Gut I, Heath S, Foglio M, Liang L, Laukens D, Mni M, Zelenika D, Van Gossum A, Rutgeerts P, Belaiche J, Lathrop M, Georges M. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. PLoS Genet. 2007; 3:e58.

144. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 2006; 314:1461-1463.

145. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004; 291:2457-2465.

146. Chan AW, Krleza-Jeric K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. CMAJ 2004; 171:735-740.

147. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. BMJ 2005; 330:753.

148. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JP. An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. Pharmacogenet Genomics 2006; 16:705-711.

149. Winker MA. Race and ethnicity in medical research: requirements meet reality. J.Law Med.Ethics 2006; 34:520-5, 480.

150. Steinberg K, Gallagher M. Assessing Genotypes in Human Genome Epidemiology Studies. In: Khoury MJ, Little J, Burke W, eds. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease New York: Oxford University Press, 2004:79-91.

151. Scuteri A, Sanna S, Chen WM, Uda M, Albai G, Strait J, Najjar S, Nagaraja R, Orru M, Usala G, Dei M, Lai S, Maschio A, Busonero F, Mulas A, Ehret GB, Fink AA, Weder AB, Cooper RS, Galan P, Chakravarti A, Schlessinger D, Cao A, Lakatta E, Abecasis GR. Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. PLoS Genet. 2007; 3:e115.

152. Begg CB. On the use of familial aggregation in population-based case probands for calculating penetrance. J Natl Cancer Inst 2002; 94:1221-1226.

153. Bourgain C, Genin E. Complex trait mapping in isolated populations: Are specific statistical methods required? Eur.J.Hum.Genet. 2005; 13:698-706.

154. Bourgain C. Comparing strategies for association mapping in samples with related individuals. BMC Genet. 2005; 6 Suppl 1:S98.

155. Gu CC, Yu K, Boerwinkle E. Measuring marker information content by the ambiguity of block boundaries observed in dense SNP data. Ann.Hum.Genet. 2007; 71:127-140.

156. Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, Balding DJ, Chokkalingam A, Dolan SM, Flanders WD, Higgins JP, McCarthy MI, McDermott DH, Page GP, Rebbeck TR, Seminara D, Khoury MJ. Assessment of cumulative evidence on genetic associations: interim guidelines. Int.J.Epidemiol. 2007 (in press);

157. Gambaro G, Anglani F, D'Angelo A. Association studies of genetic polymorphisms and complex disease. Lancet 2000; 355:308-311.

158. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ionnadis DG. Replication validity of genetic association studies. Nat Genet 2001; 29:306-309.

159. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet Med 2002; 4:45-61.

160. Tabor HK, Risch NJ, Myers RM. Opinion: candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002; 3:391-397.

161. Lancet. In search of genetic precision. Lancet 2003; 361:357-357.

162. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat.Genet. 2003; 33:177-182.

163. Gorroochurn P, Hodge SE, Heiman GA, Durner M, Greenberg DA. Non-replication of association studies: "pseudo-failures" to replicate? Genet.Med. 2007; 9:325-331.

164. Taubes G. Epidemiology faces its limits. Science 1995; 269:164-169.

165. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; 283:2008-2012.