



## Special Article

# Genome-Wide Association Studies, Field Synopses, and the Development of the Knowledge Base on Genetic Variation and Human Diseases

Muin J. Khoury, Lars Bertram, Paolo Boffetta, Adam S. Butterworth, Stephen J. Chanock, Siobhan M. Dolan, Isabel Fortier, Montserrat Garcia-Closas, Marta Gwinn, Julian P. T. Higgins, A. Cecile J. W. Janssens, James Ostell, Ryan P. Owen, Roberta A. Pagon, Timothy R. Rebbeck, Nathaniel Rothman, Jonine L. Bernstein, Paul R. Burton, Harry Campbell, Anand Chockalingam, Helena Furberg, Julian Little, Thomas R. O'Brien, Daniela Seminara, Paolo Vineis, Deborah M. Winn, Wei Yu, and John P. A. Ioannidis

Initially submitted June 6, 2008; accepted for publication April 15, 2009.

Genome-wide association studies (GWAS) have led to a rapid increase in available data on common genetic variants and phenotypes and numerous discoveries of new loci associated with susceptibility to common complex diseases. Integrating the evidence from GWAS and candidate gene studies depends on concerted efforts in data production, online publication, database development, and continuously updated data synthesis. Here the authors summarize current experience and challenges on these fronts, which were discussed at a 2008 multidisciplinary workshop sponsored by the Human Genome Epidemiology Network. Comprehensive field synopses that integrate many reported gene-disease associations have been systematically developed for several fields, including Alzheimer's disease, schizophrenia, bladder cancer, coronary heart disease, preterm birth, and DNA repair genes in various cancers. The authors summarize insights from these field synopses and discuss remaining unresolved issues—especially in the light of evidence from GWAS, for which they summarize empirical  $P$ -value and effect-size data on 223 discovered associations for binary outcomes (142 with  $P < 10^{-7}$ ). They also present a vision of collaboration that builds reliable cumulative evidence for genetic associations with common complex diseases and a transparent, distributed, authoritative knowledge base on genetic variation and human health. As a next step in the evolution of Human Genome Epidemiology reviews, the authors invite investigators to submit field synopses for possible publication in the *American Journal of Epidemiology*.

association; database; encyclopedias; epidemiologic methods; genome, human; genome-wide association study; genomics; meta-analysis

Abbreviations: dbGaP, Database on Genotypes and Phenotypes; GWAS, genome-wide association studies; HuGE, Human Genome Epidemiology; HuGENet, Human Genome Epidemiology Network; NAT2, *N*-acetyltransferase type 2; STREGA, Strengthening the Reporting of Genetic Association Studies.

**Editor's note:** This article also appears on the Web site of the Human Genome Epidemiology Network (<http://www.cdc.gov/genomics/hugenet/default.htm>).

The rapid growth in published genetic association studies (1) and the more recent success of genome-wide association

studies (GWAS) in finding disease susceptibility loci for several common diseases (2) present major challenges for knowledge synthesis and dissemination. Knowledge synthesis is needed to guide further research, drug discovery efforts (3), and translational efforts for personalized risk assessment and therapy. The recent trend toward direct-to-consumer advertising of whole-genome analysis by several

Correspondence to Dr. Muin J. Khoury, Office of Public Health Genomics, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Atlanta, Georgia 30341 (e-mail: [mukl@cdc.gov](mailto:mukl@cdc.gov)).

**Table 1.** Trends in Numbers of Published Articles on Human Genome Epidemiology, Meta-Analyses, and Genome-Wide Association Studies and Numbers of Genes Studied, by Year, 2001–2008<sup>a</sup>

Year	No. of Genes <sup>b</sup>	No. of Diseases	No. of Articles Published		
			Total	GWAS	Meta-Analyses <sup>c</sup>
2001	633	690	2,492	0	34
2002	794	855	3,196	0	45
2003	832	880	3,476	3	65
2004	1,124	1,021	4,280	0	86
2005	1,308	1,077	5,029	5	113
2006	1,502	1,109	5,364	12	155
2007	2,142	1,292	7,222	104	208
2008	3,336	1,203	7,659	134	206

Abbreviations: HuGE, Human Genome Epidemiology; GWAS, genome-wide association studies.

<sup>a</sup> Data were obtained through a HuGE Navigator query (<http://www.hugenavigator.net/>) conducted on February 14, 2009.

<sup>b</sup> Does not include the numbers of studied variants per gene, as such information is difficult to obtain. Individual genes were not counted unless they were featured in the paper.

<sup>c</sup> Includes HuGE reviews.

companies underscores the importance of a credible process for data synthesis and evaluation of the validity and utility of claims related to genetic prediction of disease risks (4–7).

In 2008, over 7,000 original articles on human genome epidemiology were published, and the annual number has been rising rapidly (8) (Table 1). Furthermore, published literature represents only a fraction of the data actually collected and analyzed. In addition, until recently, most studies have targeted 1 or a few gene variants (the candidate gene approach), but many new articles report the results of GWAS, and such studies are expected to become increasingly common. More than 300 GWAS have been published, with an accelerating pace since 2007 (8). Only a few of these studies, however, have been deposited into accessible online databases such as the Database on Genotypes and Phenotypes (dbGaP) at the National Library of Medicine (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap>), the Cancer Genetic Markers of Susceptibility database at the National Cancer Institute (<http://www.cgems.cancer.gov/>), and the Wellcome Trust Case-Control Consortium database (<http://www.wtccc.org.uk/>). This number is expected to increase under a new policy governing data-sharing for GWAS (9), although issues pertaining to type of access and confidentiality may continue to need careful consideration (10).

Despite a massive amount of primary data, the conclusions of genetic association studies are not always clear, requiring an evidence-based synthesis that takes into account the amount of evidence, the extent of replication, and protection from bias. Although approximately 1,000 systematic reviews and meta-analyses have been published since 2001, most have addressed only 1 or a few specific gene-disease associations at a time (8). Moreover, the amount of accumulated data that needs to be integrated continues to grow rapidly, with high-throughput genotyping platforms raising the challenge exponentially.

As part of ongoing efforts in this field, we report here findings and recommendations from a multidisciplinary workshop including geneticists, epidemiologists, journal editors, and bioinformatics experts that was sponsored by the Human Genome Epidemiology Network (HuGENet) and held in Atlanta, Georgia, on January 24 and 25, 2008. The meeting was convened to discuss synthesis and appraisal of cumulative evidence on genetic associations and to develop a strategy for an online encyclopedia of genetic variation and common human diseases.

## PROGRESS IN THE HuGENet ROAD MAP

HuGENet is an informal global collaboration of individuals and organizations interested in accelerating the development of the knowledge base on genetic variation and human health (11) (<http://www.cdc.gov/genomics/hugenet/default.htm>). HuGENet has developed a “road map” (12) with several components: 1) working with genetic epidemiology study platforms (primarily consortia and networks) to improve the execution and output of these groups under the rubric “Network of Networks” (13); 2) promoting the publication of results from methodologically sound genetic association studies with transparent reporting of their methods (Strengthening the Reporting of Genetic Association Studies (STREGA) (14)) and avoidance of selective reporting; 3) developing methods for synthesis and meta-analysis of the literature on genetic associations (see the *HuGENet HuGE Review Handbook, Version 1.0* (15)); and 4) developing “field synopses” with an online encyclopedia summarizing what we do and do not know about genetic associations through a systematic assessment of the cumulative evidence. Such field synopses were also called for in a 2006 *Nature Genetics* editorial (16).

## FIELD SYNOPSES: ASSESSING CUMULATIVE EVIDENCE FOR GENETIC ASSOCIATIONS

An initial meeting of the Network of Networks in 2005 led to the formation of a working group on methods for assessing cumulative evidence. A workshop organized in Venice, Italy, in 2006 (17) generated interim guidelines for grading the cumulative evidence on genetic associations based on 3 criteria: 1) the amount of evidence, 2) the extent of replication, and 3) protection from bias (18). The proposed scheme allows for 3 categories of descending credibility (A, B, C) for each of these criteria and also for a composite assessment of “strong,” “moderate,” or “weak” credibility (see Appendix Table and Ioannidis et al. (18) for more details). Briefly, an overall “strong” rating is reserved for an AAA rating, while an overall “weak” rating is reserved for associations with 1 or more C ratings. The rest are labeled as “moderate.” Note that these ratings could change over time with data accruing from additional studies. The panel also discussed issues of biologic and other experimental evidence and of the clinical importance of genetic associations. Pilot studies were planned in selected fields to assess cumulative evidence on gene-disease associations, calibrate the proposed guidelines, and integrate the findings into comprehensive field synopses.

**Table 2.** Key Characteristics of Pilot Field Synopses of Genetic Associations

	No. of Meta-Analyses	No. of Data Sets <sup>a</sup>	Threshold <sup>b</sup> for Meta-Analysis	No. of Statistically Significant Associations <sup>c</sup>	Strong <sup>d</sup> (Grade A)	World Wide Web Address
Alzheimer's disease <sup>e</sup>	228	1,072	4 data sets	53	NA	www.alzgene.org
Schizophrenia <sup>f</sup>	118	1,179	4 data sets	24	4	www.szgene.org
DNA repair genes and various cancers	241	1,087	2 independent teams	31	3	www.episat.org
Bladder cancer	36	356	3 data sets	7	1	Not yet online
Coronary heart disease	48	1,039		4	0	www.chdgene.com
Preterm birth	17	87	3 data sets	2	0	www.prebic.net
Major depression	22	131	3 data sets	6	2	Not yet online

Abbreviation: NA, not applicable.

<sup>a</sup> Total number of data sets included in the meta-analyses (not including data sets that did not undergo meta-analysis).

<sup>b</sup> Authors' prerequisite condition for conducting a meta-analysis.

<sup>c</sup> Statistically significant ( $P < 0.05$ ) by random-effects calculations on the default (per allele) analysis (for coronary heart disease, results are based on a meta-regression model and correspond to effects in the largest studies, while for DNA repair genes, both recessive and dominant models were investigated).

<sup>d</sup> Grade AAA with regard to all 3 Venice criteria (18).

<sup>e</sup> Current on February 27, 2008.

<sup>f</sup> Current on April 30, 2008.

As of 2008, pilot field synopses have been conducted for several diseases, including Alzheimer's disease, bladder cancer, schizophrenia, preterm birth, and coronary heart disease, as well as DNA repair genes and cancer phenotypes.

A field synopsis is a regularly updated snapshot of the current state of knowledge about genetic associations in a particular field of research defined by a disease (e.g., Alzheimer's disease), phenotype (e.g., body mass index), or family of genes (e.g., DNA repair genes). The ideal attributes of a field synopsis are that it: 1) is freely available; 2) uses online databases that are curated by researchers to develop regularly updated "online tables" on the volume of the evidence and the magnitude of the associations between the disease and all genetic variants investigated; 3) uses objective and transparent criteria for grading the credibility of cumulative evidence; 4) summarizes the information in peer-reviewed articles; and 5) updates information on a regular basis. The first field synopsis—the source of AlzGene, the Alzheimer's disease genetic association database—was developed by Bertram et al. (19) and published in January 2007. This was followed by the publication of a field synopsis on schizophrenia (20) and one on DNA repair genes (21), while several other synopses are currently under development or peer review.

#### EXPERIENCE WITH FIELD SYNOPSES TO DATE

At the HuGENet workshop, several teams presented findings from and experiences in developing field synopses and grading the epidemiologic evidence according to the interim Venice guidelines (18). Key features of these efforts are summarized in Table 2. All synopses included results from multiple meta-analyses involving large numbers of data sets,

except for preterm birth, where evidence is sparse. Researchers performing synopses have used different thresholds or trigger points for conducting a meta-analysis. For example, in the coronary heart disease field synopsis, investigators have considered only those associations for which at least 1 previous effort to perform a meta-analysis has been made. Data from GWAS have been incorporated in synopses on Alzheimer's disease, schizophrenia, DNA repair genes, and bladder cancer. The preterm birth field synopsis points out the need for further research on the genetic contribution to this major public health challenge.

Many associations in the Alzheimer's disease, schizophrenia, and 2 cancer-related field synopses yielded formally statistically significant results at the  $P < 0.05$  level (Table 2). Nevertheless, only a few associations met the designation of "strong" evidence according to the Venice criteria. Similarly, in several synopses, none of the probed associations attained the status of "strong" evidence. Finally, so far, field synopses have examined only 1 main phenotype, except in the case of DNA repair genes. In addition to main effects, synopses have investigated genetic effects according to different genetic models and for subgroups—for example, subgroups based on exposure, ethnic group, participant characteristics, or phenotypic subgroups. Decisions to undertake additional analyses need to be made on the basis of data availability. For example, in most field synopses, investigators were able to assess different genetic models. Often, available epidemiologic evidence may be stronger for one genetic model than for another. By contrast, there have been relatively fewer subgroup analyses based on exposures and participant characteristics, because of suboptimal reporting of these factors in genetic epidemiology studies, a deficiency that the STREGA guidance aims to address (14).

## INSIGHTS FROM CURRENT FIELD SYNOPSES

The pilot field synopses provided detailed insight about the grading process in the 3 specified areas: amount of evidence, replication, and protection from bias. They identified limitations that will help us refine the current approach.

### Amount of evidence

For amount of evidence, synopses have used a classification scheme based on the sample size of the minor genetic group (participants or alleles, depending on the genetic model). This is a simple measure that is readily available and has a close connection to power, Bayes factors, or false discovery rate (18). For candidate-gene variants, several postulated associations fail to reach grade A evidence (see Appendix Table). Currently, with large collaborative efforts stemming from GWAS and subsequent replication studies, this is likely to be less of a problem, at least for common variants with a frequency greater than 5%–10%. For variants with lower frequency, very large sample sizes may be required. Nevertheless, some consortia have the potential of reaching even sample sizes exceeding 100,000, which means more than 1,000 for the minor allele, even for variants that occur in 0.5% of the general population. For example, the international consortium on osteoporosis (Genetic Factors for Osteoporosis) funded by the European Commission includes 61 studies with 133,333 participants, and for at least 14 of these studies, investigators have already conducted or plan to conduct GWAS. We may need to revisit the criteria on amount of evidence once we have a better sense of the effect sizes regularly encountered for more rare variants.

### Replication

Field synopses have used  $I^2$  to assign grades for inconsistency (amount of heterogeneity) (i.e., A for <25%, B for 25%–50%, C for >50%) across studies (19–21). One-third to one-half of the formally significant associations have moderate or large  $I^2$  values. However,  $I^2$  often has large uncertainty when there are only a few studies (22). Moreover, qualitative epidemiologic considerations about the presence of and potential explanation for heterogeneity would need to be taken into account in judging replication. For example, the association between *N*-acetyltransferase type 2 (*NAT2*) variants and bladder cancer risk is expected to be exposure-specific; thus, heterogeneity may readily be expected between populations with different exposures (e.g., different types of tobacco in European populations vs. other populations) (23, 24).

Another consideration is whether  $I^2$  reflects heterogeneity of estimates around the null value or heterogeneity in the magnitude of association. The former would question the presence of an association, whereas the latter would question the strength of the association. For instance, even for a consistent association such as the glutathione *S*-transferase M1 null genotype and bladder cancer risk, there is some evidence for heterogeneity in the magnitude of the association across studies (24). However, such epidemiologic in-

sight must be considered with caution, to avoid introducing subjective, speculative processes in the grading. At a minimum, considerations for upgrading or downgrading should be explicit. It may be reasonable to grade as A on this criterion associations with moderate or high heterogeneity with an extensive replication record. This replication includes a  $P$  value for the summary effect (excluding the discovery data set) of  $P < 10^{-7}$ , even in random-effect models that account for between-study heterogeneity or have a false-positive report probability rate less than 10% or a Bayes factor less than  $10^{-5}$ .

For example, the apparent heterogeneity in the effect of *NAT2* slow acetylation on bladder cancer risk can be explained by differences in the pattern of tobacco smoking across study populations (24). However, the presence of heterogeneity would reflect even in these cases the possibility that, bias set aside, one would need to identify the sources of heterogeneity in subsequent studies. These could include not only differential effects under different exposures but also the possibility that the association is with a correlated phenotype and not the one tested (e.g., the fat mass and obesity-associated (*FTO*) gene, diabetes, and obesity) (25), the impact of the different ascertainment schemes used in different studies (26), genotype misclassification (especially in isolated candidate gene studies), or a marker polymorphism that is in variable linkage disequilibrium with the causative variant across the populations (27). The latter scenario could become common in associations that emerge out of “agnostic” GWAS, where it is unlikely that the causal variant will be directly identified. In the setting of GWAS, it is easy to check whether linkage disequilibrium structures are different in different populations; in the presence of similar linkage disequilibrium structure, a cause of heterogeneity can be quickly excluded. It has been demonstrated that beyond a given threshold of inconsistency, no matter how large the studies we conduct, we may never have enough power to replicate an association (nonreplicability threshold) (28).

Another issue is the ability of the cumulative evidence to exclude an association based on lack of replication. It is notable that the Venice criteria include, under “replication C,” also the possibility of “no association and failed replication,” based on traditional nonsignificant results for the meta-analysis. Minute effects can never be excluded, and in fact, in GWAS, many true associations yield modest results that do not cross genome-wide association  $P$ -value thresholds. Many true findings do not rise to the top of the single nucleotide polymorphism  $P$ -value ranks in phase 1 of a genome-wide association study (29). Despite extremely large sample sizes and cumulative meta-analyses of many GWAS, many associations may remain undiscovered and/or inconclusive. The Venice criteria should not be used to conclude that there is strong evidence for a null association.

### Protection from bias

A research finding cannot reach sufficient credibility (>50%) unless the probability of a false-positive association due to bias is less than the prestudy odds of an association's

**Table 3.** Some Checks for Retrospective Meta-Analyses in Field Synopses of Genetic Associations

General checks for the occurrence of or susceptibility to potential problems <sup>a</sup>
Small effect size (e.g., odds ratio less than 1.15-fold from the null value)
Association lost with exclusion of first study
Association lost with exclusion of HWE-violating studies or with adjustment for HWE
Evidence for small-study effect in an asymmetry regression test with proper type I error (52)
Evidence for excess of single studies with formally statistically significant results (53)
Topic- or subject-specific checks: consider whether they are problems
Unclear/misclassified phenotypes with possible differential misclassification against genotyping
Differential misclassification of genotyping against phenotypes
Major concerns for population stratification (need to justify for affecting odds ratio greater than 1.15-fold; not invoked to date)
Any other reason (case-by-case basis) that would render the evidence for association highly questionable

Abbreviation: HWE, Hardy-Weinberg equilibrium.

<sup>a</sup> All general checks are likely to have only modest, imperfect sensitivity and specificity for detecting problems. In particular, for effect size, a small effect size may very well reflect a true association, since many genetic associations have small effect sizes. However, if this effect has been documented in a retrospective meta-analysis that is susceptible to publication and other reporting biases, it also needs to be replicated in a prospective setting where such biases cannot operate before high credibility can be attributed to it.

being true (30). The Venice criteria include an extensive checklist for sources of biases in different settings. The checklist has different considerations depending on whether the evidence comes from retrospective meta-analyses of published data or prospective GWAS and replication studies from collaborative consortia with harmonization of data collection and analysis.

Bias checks that have been adopted in these synopses for retrospective meta-analysis include automated checks that can be readily applied to all meta-analyses of published data. These are shown in Table 3, along with a list of issues that need to be considered. General checks (that can be applied automatically to all fields) have the advantage of being objective and unambiguous, but they cannot provide definitive proof for the presence or absence of bias. For instance, a small effect size (e.g., odds ratio < 1.15) could be explained by bias, but many of the confirmed associations between single nucleotide polymorphisms and chronic diseases are of this order of magnitude. Therefore, small effect sizes, if seen consistently across many studies and with no evidence for publication bias, should not be automatically penalized. For prospective evidence, such as data accumulated from 1 or more GWAS with prospective replication across several teams in a consortium or prospective meta-analysis of many GWAS from collaborative studies (31), the considerations are quite different. Here, the small magnitude of the effect size should not be invoked as evidence of lack of protection from bias, and similarly small-study effect bias or an excess of single studies with significant findings is not an issue here, provided that there is no selective reporting of results (there is no reason for such selective reporting in a consortium).

For example, in the schizophrenia synopsis (20), of the 24 associations with nominal statistical significance, 9 were graded as “A” and 15 as “C” for “protection from bias.” The main reasons for low grades were a small summary odds ratio (odds ratio < 1.15) in what are retrospective meta-analyses of published data ( $n = 6$  associations) and

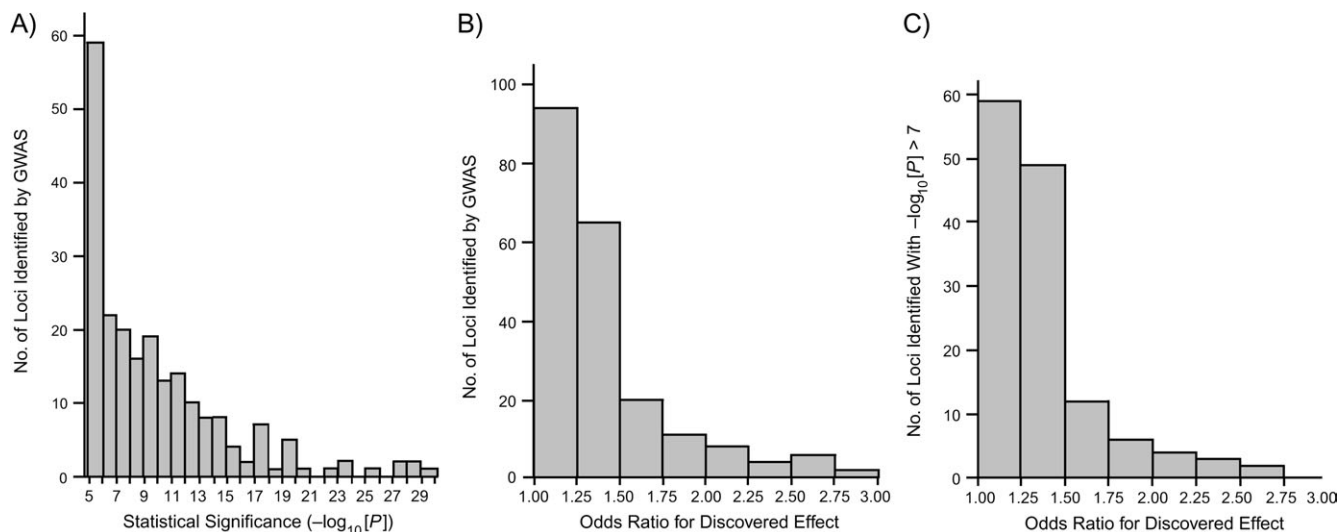
loss of significance after exclusion of the initial study ( $n = 6$ ). Less common reasons were loss of significance after exclusion of studies that violated Hardy-Weinberg equilibrium and significant differences in effect between small and larger studies.

## ISSUES TO CONSIDER FOR MOVING FORWARD

### Defining thresholds for evaluating credibility

The threshold for considering an association for further assessment must be defined in each synopsis, but it may be difficult to reach full consensus on this issue. Given that current synopses have used a large amount of evidence from candidate gene studies, most have considered for grading all probed associations that pass very lenient levels of statistical significance in meta-analysis (typically,  $P < 0.05$  inferred from random-effects calculations). However, experience to date indicates that associations with a grade of A for the amount of evidence but  $P$  values just below 0.05 have either very small effects (and get a C for protection from bias if a retrospective meta-analysis) or moderate/large heterogeneity (and thus get a B or C for replication consistency). Even for such associations that stem from the candidate gene era, it is uncommon to get a rating of “strong” epidemiologic evidence grading (AAA), unless the  $P$  value for the summary effect is substantially lower. Associations that arise out of GWAS require an even more demanding threshold. Thresholds may be set either on the basis of  $P$ -value criteria for genome-wide significance or using Bayesian approaches, of which there are several variants (32–35).

In view of the potential multiplicity of phenotypes examined and analyses performed, some authors believe that the rigorous criteria for statistical significance used in GWAS should be applied to candidate gene-derived associations. If so,  $P$  values of  $10^{-7}$  or lower would be required for a locus to be considered “confirmed” (36, 37). Figure 1, part A, shows the distribution of  $P$  values of the loci identified by



**Figure 1.** A) Levels of statistical significance for associations of genetic loci with  $P$  values of  $10^{-5}$  or lower identified through genome-wide association studies (GWAS) and entered in the National Human Genome Research Institute catalog of GWAS as of October 14, 2008 (38, 39); data are limited to those loci that have binary outcome phenotypes ( $n = 223$ ). For details on selection of loci in the catalog, see Hindorff et al. (38) and Manolio et al. (39). B) Odds ratios (per allele) for the 223 associations. C) Odds ratios for the 142 of the 233 associations that had  $P$  values less than  $10^{-7}$ . Not shown are the 5, 13, and 7 outliers that had values outside of the depicted range in the 3 panels, respectively.

GWAS for binary outcome phenotypes and which have been included in the National Human Genome Research Institute GWAS catalog as of October 14, 2008 (38, 39). Of the 466 entries in the catalog, after excluding those pertaining to studies that did not reach any hits with  $P < 10^{-5}$  and those that had nonbinary outcomes, 223 loci are included here. As shown, fewer than two-thirds of them (142/223) have a  $P$  value less than  $10^{-7}$  and only 39% (87/223) have a  $P$  value less than  $10^{-10}$ . When several studies and data sets are combined in genome-wide investigations, typically researchers have used pooled, stratified, or simple fixed-effects analyses; random-effects or other approaches that also take into account the heterogeneity between data sets would have often yielded even more conservative  $P$  values (40). This suggests that the majority of signals emerging from current GWAS and early replication efforts do not yet cross stringent levels of “genome-wide significance.” This further highlights the need to include far more data from additional GWAS and replication data sets, and this can be routinely accomplished in the setting of field synopses collating all of this information.

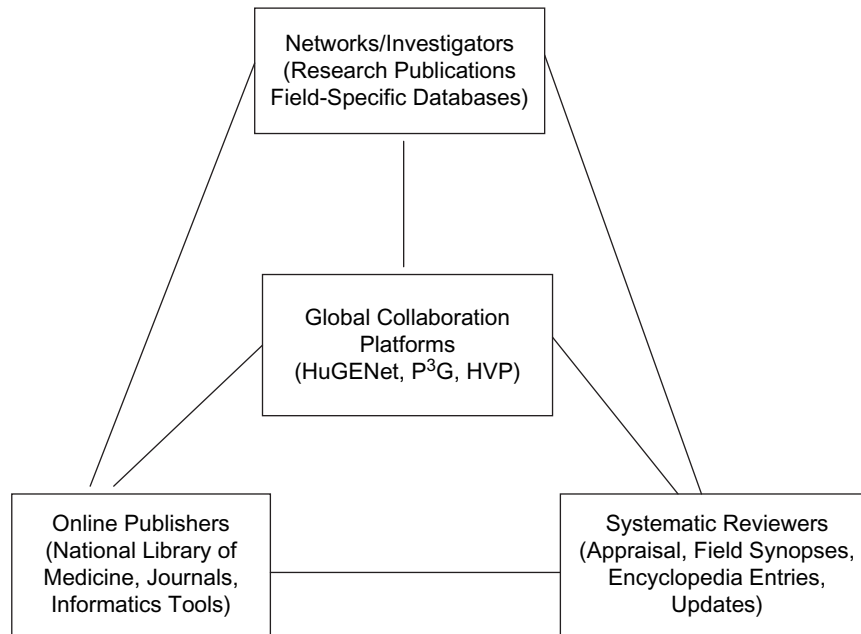
Bayesian approaches offer the advantage of allowing different prior probabilities for an association’s being present based on external evidence (thus bridging agnostic and candidate approaches) (32–35). These methods also allow consideration of the impact of different assumptions about the genetic effect sizes. Empirical evidence from GWAS can offer insight about typical discovered effects. Figure 1, part B, shows the distribution of odds ratios (typically per allele, as reported in the National Human Genome Research Institute catalog) among the 223 loci discovered in GWAS. As shown, the median effect corresponds to an odds ratio of 1.28, and the same median is seen for the 142 associations

with  $P < 10^{-7}$  (Figure 1, part C). These estimates may be inflated in comparison with the true effects, because of the “winner’s curse” phenomenon (inflation of effects selected based on significance thresholds) (41, 42). A median true odds ratio of 1.1–1.2 is therefore reasonable for these associations, and some effects may be even smaller, while exceptions of large odds ratios are probably uncommon. Nevertheless, one should acknowledge that the effect of the causal factor that is in the neighborhood of the tagging polymorphism may be larger, and we cannot yet exclude the possibility of considerably larger odds ratios for low-frequency variants (43). Such variants were not assessed in the first wave of GWAS, but they are being increasingly targeted in current and future efforts (44, 45).

As more synopses accrue, we can examine the stability of the Venice grading for various associations. This will help us understand whether some types of associations can change from having weak credibility to having strong credibility (and vice versa). As is described below, gathering empirical evidence into field synopsis databases will allow greater insight in the assessment of cumulative evidence on genetic associations.

### Defining conglomerate evidence

It is already established practice for hypotheses about specific postulated associations to be tested using data from combinations of prospective consortia analyses stemming from GWAS and their meta-analyses and replication studies; possibly several consortia working on the same disease and phenotypes; additional scattered studies by teams that are not included in any of the consortia; and even retrospective meta-analyses encompassing some/many/all of these



**Figure 2.** A vision for collaboration among disease- and gene-specific investigators, systematic reviewers, and online publishers. HuGENet, Human Genome Epidemiology Network; HVP, Human Variome Project; P<sup>3</sup>G, Public Population Project in Genomics.

sources of data. Such “conglomerate evidence” from various sources of data may appear in various time sequences. The Venice criteria suggested that one should consider the highest possible level of evidence when data come from disparate sources. Perhaps the best currently available source is a well-designed prospective consortium analysis including several teams that have performed GWAS and replications. The results of such an analysis should have a much greater weight than the results of scattered smaller studies. If the consortium evidence results in “strong” evidence, it would not be reasonable to underrate this evidence because of a few small, scattered, inconclusive studies. However, the challenge will become more serious when many consortia with 1 or more genome-wide platforms are available, and when the scattered or retrospectively meta-analyzed data are much larger in amount than the original consortium-level data on which the reported association was based. Dealing efficiently with this situation requires transparent and comprehensive availability of the evidence from these diverse studies, as we discuss below.

#### GLOBAL COLLABORATION: FROM DATA TO KNOWLEDGE

After reviewing pilot field synopses, participants in the HuGENet workshop discussed how to link emerging data on genetic associations with other sources of information on the biology of genes and gene-disease relations. Clearly, the advent of GWAS in large-scale collaborative studies involving networks and consortia is a crucial first step towards the generation of large-scale data sets. Furthermore, the deposition of these data in accessible public databases can help

to address the problem of publication bias commonly seen in candidate gene association studies. Nevertheless, additional efforts are needed to transform data into a knowledge base. Systematic reviews and meta-analyses represent a crucial step in building the knowledge base on genetic variation and human health. Such efforts need to be transparent and their results made available in online databases and publications. The willingness of journal editors to contribute to these efforts is critical, as investigators and systematic reviewers struggle to gain academic recognition for their work, which is often part of multinational, multiple-investigator studies. Finally, the National Library of Medicine (<http://www.ncbi.nlm.nih.gov/>) plays a leading role in linking genetic association studies with other existing databases on gene sequences, products, and linkages to disease processes.

At the HuGENet workshop, a vision emerged of collaboration to create a sustainable, credible knowledge base on genetic variation and human diseases. As Figure 2 shows, the collaboration involves research investigators, systematic reviewers, online publishers, and database developers with variable degrees of overlap among the groups. For example, investigators who are part of research consortia have their own informatics tools and databases, and they can conduct systematic reviews of their own field based on their own data or also including data from teams external to the consortium. In addition, other reviewers could contribute to these efforts, as evidenced by many previous efforts in meta-analyses and Human Genome Epidemiology (HuGE) reviews. Figure 2 shows the flow from generation of new data to systematic appraisal and synthesis and to online dissemination via journals and databases.

A successful example of collaboration already exists in the field of type 2 diabetes. Investigators from diverse consortia have combined efforts to conduct comprehensive meta-analyses of all GWAS and replication studies. A first meta-analysis combined 3 GWAS with a total of over 10,000 samples; this was followed by a second stage of replication of the most interesting signals in over 22,000 independent samples and a subsequent third stage of replication on over 57,000 samples, with data being combined by means of formal meta-analysis methods (46). In another example, the GIANT [Genetic Investigation of Anthropometric Traits] consortium has already published results of a meta-analysis including data from 15 GWAS, followed by replication in another 14 cohorts (47). Similar meta-analyses are being designed and carried out by collaborating consortia in several other fields—for example, the Psychiatric GWAS Consortium, which is conducting meta-analyses within and between 5 psychiatric disorders (<https://pgc.unc.edu/faqs.html>).

Several global collaborations focused on genotype-phenotype correlations can help support fields where large-scale studies are still in the making. For example, HuGENet sponsors the HuGE Navigator (8) (<http://www.hugenavigator.net/>), a knowledge base with online tools for capturing and organizing the most up-to-date information on genetic associations and other related information. The Human Variome Project is focused on the production and synthesis of gene- and gene-variant-centered databases with linked phenotypic outcomes (48). The Public Population Project in Genomics (<http://www.p3gconsortium.org/>) aims to harmonize data collected from large-scale cohort studies and biobanks around the world. Cross-links among HuGENet, the Public Population Project in Genomics, the Human Variome Project, and other groups are crucial to convene and facilitate collective efforts in developing the knowledge base on genetic variation and human diseases. Efforts in coordinating these global collaborations are already under way through cross-linking of these enterprises. For example, the Public Population Project in Genomics has an international working group in epidemiology and biostatistics that is closely related to the HuGENet movement. Another, more specialized online knowledge base development effort that can be synergistic is the Pharmacogenomics Knowledge Base (49). In addition, GeneReviews are expert-authored, peer-reviewed disease descriptions focused on the use of genetic testing in the diagnosis, management, and genetic counseling of patients and their families. GeneReviews are part of the GeneTests Web site (<http://www.genetests.org>), which also includes international directories of genetics clinics and genetics laboratories (50). Finally, it is important for epidemiologic efforts to be linked with biologic efforts, including experimental work, assessment of endophenotypes, and functional studies in different model systems.

#### **SCHIZOPHRENIA: FIELD SYNOPSIS AND EXAMPLE OF DEVELOPMENT OF A KNOWLEDGE BASE**

As an example of the collaboration among primary investigators, systematic reviewers, and online publishers,

Bertram et al. (19) provide a model approach to a distributed knowledge base of genetic variants that features collaboration among the 3 groups outlined above. They have synthesized primary research on genetic associations in schizophrenia, and they developed a regularly updated online database, SZGene ([www.szgene.org](http://www.szgene.org)), that collects and curates published results in this area. A peer-reviewed field synopsis summarizes the cumulative evidence and evaluates it according to the Venice criteria. The field synopsis is regularly updated online with updated cumulative meta-analyses. Bertram et al. have developed similar resources for Alzheimer's disease ([www.alzgene.org](http://www.alzgene.org)).

The HuGE Navigator Web site (<http://www.hugenavigator.net/>) serves to link field-specific efforts like SZGene with other online databases through the HuGEpedia. The HuGEpedia can be accessed by using either a phenotype (Phenopedia) or a gene (Genopedia) as the starting point. For example, searching the Phenopedia for schizophrenia leads users to a page that provides an up-to-date summary of genes studied for a possible association with schizophrenia, links to abstracts of the original publications in PubMed, meta-analyses and HuGE reviews, and abstracted meta-tables. HuGE Navigator can also be searched to locate investigators in the field and to display geographic and temporal trends in the published literature. Finally, HuGE Navigator attempts to identify and link to all published GWAS in the field, as well as to data sets deposited and available through the National Center for Biotechnology Information's dbGaP. Although HuGE Navigator is not a comprehensive data repository, it serves as a first stop for orientation and links to more authoritative data sources and field synopses. The highest level of data integration in this example occurs through links with National Center for Biotechnology Information databases (such as PubMed, Entrez Gene, and dbGaP). The National Center for Biotechnology Information's online book, *Genes and Disease* (51), could also expand to accommodate the most current synopses in individual fields.

#### **CONCLUDING REMARKS: INVITATION TO SUBMIT FIELD SYNOPSES FOR PUBLICATION**

This is a crucial time in human genomics research, when advances in genome-wide analysis platforms coupled with declining costs are producing an unprecedented outpouring of replicated genetic associations with common diseases. To make the most of the research enterprise and to promote reliable and timely knowledge synthesis, the multidisciplinary working group offers the following recommendations.

First, data from GWAS should be made available for interested researchers in order to avoid selective positive reporting of spurious associations and to facilitate meta-analyses of particular associations. Involvement of the primary investigators from the GWAS in collaborative projects and meta-analyses should be encouraged. There is a risk of errors' and misconceptions' being introduced if the primary investigators who are intimately familiar with the data are not involved. Second, researchers and research networks should develop field synopses that use meta-analysis to integrate published and unpublished data and evaluate the



cumulative evidence. The Venice guidelines offer interim guidance, and further empirical research is needed to assess the stability and implementation of these guidelines. Third, we encourage the development of field-specific databases, such as the SZGene database discussed above. Fourth, we encourage journal editors to publish field synopses with regular updates, as called for by *Nature Genetics* in 2006 (16). Fifth, we recommend that journals and online publishers develop and make widely available databases that include standardized and systematically collected information from original research for research synthesis. The HuGE Navigator is one approach presented here, but others could emerge in the future. The rapidity of data accumulation necessitates such a systematic approach as a starting point for evaluating the gaps in our knowledge base. To succeed, these efforts depend on collaboration fueled by the availability of funding not only for generating original research data but also for efforts in research synthesis and dissemination. Finally, we need to ensure that the epidemiologic research synthesis discussed here is accompanied by critical appraisal and synthesis of biologic research. The combination of epidemiology and biology is crucial to enhance the credibility of genetic associations and to accelerate their applications in clinical medicine and population health.

We invite prospective authors to submit field synopses to the *American Journal of Epidemiology* using the criteria laid out above with an assessment of cumulative evidence on genetic associations. This is a crucial next step beyond HuGE reviews and meta-analyses of the association between 1 gene and 1 disease, into comprehensive development of the knowledge base pertaining to specific fields. Additional instructions on preparing field synopses can be found on the HuGENet Web site (<http://www.cdc.gov/genomics/hugenet/default.htm>).

## ACKNOWLEDGMENTS

Author affiliations: Office of Public Health Genomics, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia (Muin J. Khoury, Marta Gwinn, Wei Yu); Genetics and Aging Research Unit, Massachusetts General Hospital, Charlestown, Massachusetts (Lars Bertram); International Agency for Research on Cancer, Lyon, France (Paolo Boffetta); Molecular Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom (Adam S. Butterworth, Julian P. T. Higgins); Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland (Stephen J. Chanock, Montserrat Garcia-Closas, Nathaniel Rothman, Thomas R. O'Brien); Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York (Siobhan M. Dolan); Quebec Innovation Center, McGill University, Montreal, Quebec, Canada (Isabel Fortier); MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom (Julian P. T. Higgins); Erasmus University Medical Center, Rotterdam, The Netherlands (A. Cecile J. W. Janssens); National Center for Biotechnology Informa-

tion, National Library of Medicine, Bethesda, Maryland (James Ostell); Stanford University Medical Center, Stanford, California (Ryan P. Owen); School of Medicine, University of Washington, Seattle, Washington (Roberta A. Pagon); School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania (Timothy R. Rebbeck); Sloan-Kettering Cancer Center, New York, New York (Jonine L. Bernstein); Department of Health Sciences, University of Leicester, Leicester, United Kingdom (Paul R. Burton); Center for Population Health Sciences, University of Edinburgh, Edinburgh, United Kingdom (Harry Campbell); Division of Epidemiology, University of California at Berkeley, Berkeley, California (Anand Chockalingam); Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Helena Furberg); Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada (Julian Little); Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland (Daniela Seminara, Deborah M. Winn); Division of Epidemiology, Imperial College London, London, United Kingdom (Paolo Vineis); Department of Hygiene and Epidemiology, School of Medicine, University of Ioannina, Ioannina, Greece (John P. A. Ioannidis); Tufts Clinical and Translational Science Institute, Tufts University, Boston, Massachusetts (John P. A. Ioannidis); and Center for Genetic Epidemiology and Modeling, Tufts Medical Center, Boston, Massachusetts (John P. A. Ioannidis).

Scientific support for Dr. John P. A. Ioannidis was provided through the Tufts Clinical and Translational Science Institute (Tufts University, Boston, Massachusetts) under funding from the US National Institutes of Health/National Center for Research Resources (UL1 RR025752). Drs. Isabel Fortier and Paul R. Burton were supported by the Public Population Project in Genomics (Montreal, Quebec, Canada), which is funded by Genome Canada and Genome Quebec, and PHOEBE [Promoting Harmonisation of Epidemiological Biobanks in Europe], which is funded by the European Union under the Sixth Framework Programme.

The authors thank Drs. Myles Axton, Teri Manolio, and Cynthia Morton.

The findings and conclusions presented in this report are those of the authors and do not necessarily reflect the view of the US Department of Health and Human Services or the official position or policies of the Tufts Clinical and Translational Science Institute.

Conflict of interest: none declared.

## REFERENCES

1. Lin B, Clyne M, Walsh M, et al. Tracking the epidemiology of human genes in the literature: the HuGE published literature database. *Am J Epidemiol*. 2006;164(1):1-4.
2. Topol EJ, Murray SS, Frazer KA. The genomics gold rush. *JAMA*. 2007;298(2):218-221.
3. Kingsmore SF, Lindquist IE, Mudge J, et al. Genome-wide association studies: progress and potential for drug discovery and development. *Nat Rev Drug Discov*. 2008;7(3):221-230.

4. Hunter DJ, Khoury MJ, Drazen JM. Letting the genome out of the bottle—will we get our wish? *N Engl J Med*. 2008; 358(2):105–107.
5. Risky business [editorial]. *Nat Genet*. 2007;39(12):1415.
6. Positively disruptive [editorial]. *Nat Genet*. 2008;40(2):119.
7. NCI-NHGRI Working Group on Replication in Association Studies; Chanock SJ, Manolio T. Replicating genotype-phenotype associations. *Nature*. 2007;447(7145):655–660.
8. Yu W, Gwinn M, Clyne M, et al. A navigator for human genome epidemiology. *Nat Genet*. 2008;40(2):124–125.
9. National Institutes of Health. *Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS)*. (Notice no. NOT-OD-07-088). Bethesda, MD: National Institutes of Health; 2007. (<http://www.grants.nih.gov/grants/guide/notice-files/NOT-OD-07-088.html>). (Accessed April 19, 2008).
10. Homer N, Szelling S, Redman M, et al. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays [electronic article]. *PLoS Genet*. 2008;4(8):e1000167.
11. Khoury MJ, Dorman JS. The Human Genome Epidemiology Network. *Am J Epidemiol*. 1998;148(1):1–3.
12. Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet*. 2006; 38(1):3–5.
13. Ioannidis JP, Bernstein J, Boffetta P, et al. A network of investigator networks in human genome epidemiology. *Am J Epidemiol*. 2005;162(4):302–304.
14. Little J, Higgins JP, Ioannidis JP, et al. Strengthening the Reporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. *Ann Intern Med*. 2009;150(3): 206–215.
15. Little J, Higgins JP, eds. *The HuGENet™ HuGE Review Handbook, Version 1.0*. Ottawa, Ontario, Canada: HuGENet Canada Coordinating Centre; 2006. (<http://www.hugenet.ca>). (Accessed March 10, 2008).
16. Embracing risk [editorial]. *Nat Genet*. 2006;38(1):1.
17. Office of Public Health Genomics, Centers for Disease Control and Prevention. *HuGENet™ Workshop on the Assessment of Cumulative Evidence on Genetic Associations: International Workshop held in Venice (Italy) on 9–10 November 2006*. Atlanta, GA: Centers for Disease Control and Prevention; 2006. ([http://www.cdc.gov/genomics/hugenet/hugewkshp\\_nov06.htm](http://www.cdc.gov/genomics/hugenet/hugewkshp_nov06.htm)). (Accessed April 15, 2008).
18. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol*. 2008;37(1):120–132.
19. Bertram L, McQueen MB, Mullin K, et al. Systematic meta-analyses of Alzheimer genetic association studies: the AlzGene database. *Nat Genet*. 2007;39(1):17–23.
20. Allen NC, Bagade S, McQueen MB, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet*. 2008;40(7): 827–834.
21. Vineis P, Manuguerra M, Kavvoura FK, et al. A field synopsis on low-penetrance variants in DNA repair genes and cancer susceptibility. *J Natl Cancer Inst*. 2009;101(1): 24–36.
22. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007; 335(7626):914–916.
23. García-Closas M, Malats N, Silverman D, et al. NAT2 slow acetylation and GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet*. 2005;366(9486):649–659.
24. Rothman N, Garcia-Closas M, Hein DW. Commentary: reflections on G.M. Lower and colleagues' 1979 study associating slow acetylator phenotype with urinary bladder cancer: meta-analysis, historical refinements of the hypothesis, and lessons learned. *Int J Epidemiol*. 2007;36(1):23–28.
25. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007; 316(5826):889–894.
26. Burton PR, Palmer LJ, Jacobs K, et al. Ascertainment adjustment: where does it take us? *Am J Hum Genet*. 2000;67(6): 1505–1514.
27. Ioannidis JP. Non-replication and inconsistency in the genome-wide association setting. *Hum Hered*. 2007;64(4):203–213.
28. Moonesinghe R, Khoury MJ, Liu T, et al. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proc Natl Acad Sci U S A*. 2008;105(2):617–622.
29. Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet*. 2008;40(3):310–315.
30. Ioannidis JP. Why most published research findings are false [electronic article]. *PLoS Med*. 2005;2(8):e124.
31. Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. 2007;316(5829):1336–1341.
32. Wakefield J. Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet*. 2007;81(2):208–227.
33. Wakefield J. Bayes factors for genome-wide association studies: comparison with *P*-values. *Genet Epidemiol*. 2008; 33(1):79–86.
34. Wacholder S, Chanock S, Garcia-Closas M, et al. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*. 2004;96(6):434–442.
35. Ioannidis JP. Calibration of credibility of agnostic genome-wide associations assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *Am J Med Genet B Neuropsychiatr Genet*. 2008;147B(6): 964–972.
36. Hoggart CJ, Clark TG, De Iorio M, et al. Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol*. 2008;32(2):179–185.
37. Pe'er I, Yelensky R, Altshuler D, et al. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol*. 2008;32(4): 381–385.
38. Hindorf LA, Junkins HA, Manolio TA. *A Catalog of Published Genome-Wide Association Studies*. Bethesda, MD: Office of Population Genomics, National Human Genome Research Institute; 2008. (<http://www.genome.gov/26525384>). (Accessed October 14, 2008).
39. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest*. 2008;118(5):1590–1605.
40. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations [electronic article]. *PLoS ONE*. 2007;2(9):e841.
41. Zollner S, Pritchard JK. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am J Hum Genet*. 2007;80(4):605–615.
42. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology*. 2008;19(5):640–648.
43. Wright AF, Charlesworth B, Rudan I, et al. A polygenic basis for late-onset disease. *Trends Genet*. 2002;19(2):97–106.

44. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet.* 2008;40(6):695–701.
45. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science.* 2008;320(5875):539–543.
46. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies several additional susceptibility loci for type 2 diabetes. *Nat Genet.* 2008;40(5):638–645.
47. Willer CJ, Speliotes EK, Loos RJ, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet.* 2009;41(1):25–34.
48. Cotton RG, Appelbe W, Auerbach AD, et al. Recommendations of the 2006 Human Variome Project meeting. *Nat Genet.* 2007;39(4):433–436.
49. Klein TE, Chang JT, Cho MK, et al. Integrating genotype and phenotype information: an overview of the PharmGKB project. Pharmacogenetics Research Network and Knowledge Base. *Pharmacogenomics J.* 2001;1(3):167–170.
50. Pagon RA. GeneTests: an online genetic information resource for healthcare providers. *J Med Libr Assoc.* 2006;94(3):343–348.
51. National Center for Biotechnology Information, National Institutes of Health. *Genes and Disease* [electronic book]. Bethesda, MD: National Center for Biotechnology Information; 2008. (<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View.ShowSection&rid=gnd.preface.91>). (Accessed February 19, 2008).
52. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med.* 2006;25(20):3443–3457.
53. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials.* 2007;4(3):245–253.

**Appendix Table.** Considerations for Epidemiologic Credibility in the Assessment of Cumulative Evidence on Genetic Associations<sup>a</sup>

Criteria and Categories	Proposed Operationalization
Amount of evidence	
A: Large-scale evidence	Thresholds may be defined on the basis of sample size, power, or false-discovery rate considerations. The frequency of the genetic variant of interest should be accounted for. As a simple rule, we suggest that category A require a sample size of more than 1,000 (total number in cases and controls, assuming a 1:1 ratio) evaluated in the least common genetic group of interest; that B correspond to a sample size of 100–1,000 evaluated in this group; and that C correspond to a sample size of less than 100 evaluated in this group (see “Discussion” section in the text and Table 2 for further elaboration). <sup>b</sup>
B: Moderate amount of evidence	
C: Little evidence	
Replication	
A: Extensive replication including at least 1 well-conducted meta-analysis with little between-study inconsistency	Between-study inconsistency entails statistical considerations (e.g., defined by metrics such as $I^2$ , where values of 50% and above are considered large and values of 25%–50% are considered moderate inconsistency) and also epidemiologic considerations for the similarity/standardization or at least harmonization of phenotyping, genotyping, and analytical models across studies. See “Discussion” section in the text for the threshold (statistical or other) required for claiming replication under different circumstances (e.g., with or without inclusion of the discovery data in situations with massive testing of polymorphisms).
B: Well-conducted meta-analysis with some methodological limitations or moderate between-study inconsistency	
C: No association; no independent replication; failed replication; scattered studies; flawed meta-analysis or large inconsistency	
Protection from bias	
A: Bias, if at all present, could affect the magnitude but probably not the presence of the association	A prerequisite for A is that the bias due to phenotype measurement, genotype measurement, confounding (population stratification), and selective reporting (for meta-analyses) can be appraised as not being high (as shown in detail in Table 3)—plus, there is no other demonstrable bias in any other aspect of the design, analysis, or accumulation of the evidence that could invalidate the presence of the proposed association. In category B, although no strong biases are visible, there is no such assurance that major sources of bias have been minimized or accounted for, because information is missing on how phenotyping, genotyping, and confounding have been handled. Given that occult bias can never be ruled out completely, note that even in category A, we use the qualifier “probably.”
B: No obvious bias that may affect the presence of the association, but there is considerable missing information on the generation of evidence	
C: Considerable potential for or demonstrable bias that can affect even the presence or absence of the association	

<sup>a</sup> Based on the Venice criteria (18).

<sup>b</sup> For example, if the association pertains to the presence of homozygosity for a common variant and if the frequency of homozygosity is 3%, then category A under “Amount of evidence” requires over 30,000 subjects and category B between 3,000 and 30,000. The sample size refers to subjects when genotype contrasts are used and to alleles when alleles are contrasted.