

Position Statement of the International Genetic Epidemiology Society  
in response to

*“Request for Information (RFI): Proposed Policy for Sharing of Data obtained in NIH supported or conducted Genome-Wide Association Studies (GWAS)”*

The International Genetic Epidemiology Society (IGES) has a long history of advocating the sharing of scientific knowledge, and when appropriate, the sharing of data from large genetic epidemiological studies, where data can range from subject-specific data to summary data. The IGES also recognizes that the potential advantages of such data sharing must be balanced against various scientific concerns, as well as the critical need to protect the confidentiality of the participants in the studies for which data are shared. **A consensus has been growing among geneticists that genome-wide association study (GWAS) genotype data can never be completely de-identified, since genotypes are themselves identifiers.**

Many IGES members have expressed serious concerns about the proposed policy by the NIH that requires deposition of genotype and phenotype data from GWAS funded or supported by the NIH immediately into US Government databases (Notice Number: NOT-OD-06-094). Deposition of data into such U.S. Government databases carries the risk that U.S. Federal law enforcement agencies (such as the FBI, CIA, Homeland Security) can legally search these databases without a court ordered subpoena, whereas a subpoena is required for these agencies to obtain access to data stored in non-Federal databases. In addition, the proposed policy does not address concerns about storing biometric identifiers of non-citizens in U.S. Federal databases, which may deleteriously affect international collaborations. **We recommend that original study investigators should be allowed to maintain the data in non-Federal databases to minimize these risks.**

Many IGES members are also concerned with the plan that data requestors should only have to sign a certification that they will protect the data, without any IRB (Human Subjects Protection Committee or Ethics Board) oversight at their own institution. Under the current NIH broad data sharing policy, the original study investigators and/or their IRB have the ability to require that the data requestors must present evidence that they have completed human subjects training (as is required of all NIH grantees who will take part in grants that involve human subjects data) and must also obtain some level of IRB oversight and approval of their proposed project. This extra level of IRB oversight is particularly important when data are made available to investigators who have no contact with the original investigators and who may not be aware of the special issues concerning protection of genetic data. **Therefore, we recommend that original study investigators should retain the option to require that data requestors for GWAS genotypic data must obtain an IRB approval, to ensure the human subjects protections guaranteed by study investigators of most existing studies.**

There are additional concerns that bias and non-representativeness may be introduced into GWAS studies, making results suspect, and limiting generalizability of results to multiple populations if this new policy is implemented. It is expected that some biases will occur if very important collections of data are no longer eligible for GWAS if the proposed rules are adopted. Other biases can occur because certain populations may be underrepresented in broadly released current data collections if many of their members refuse to re-consent to have their data deposited in federal databases. Future data collection may have similar problems with diversity if potential study subjects refuse to participate because of the real risk that they might be identified.

In contrast to the proposed new NIH policy, the IGES believes that the current NIH data sharing policy for studies with direct costs of \$500,000 or more per year is both adequate and fair. This existing policy mandates that broad data sharing should take place, but it allows some initial exclusive use of the data by the original investigators and it further allows the original investigators to more carefully protect the human subjects who participate in their studies. Similar policies have been followed by the NIMH Genetics Initiative for the past ten years with no documented adverse consequences. The broad data sharing policies followed by the Framingham

Heart Study, Cancer Family Registries and Family Blood Pressure Program are appropriate models for broad data sharing that adequately protect the interests of study subjects and investigators while advancing scientific goals. **The IGES recommends that the existing broad data sharing policy should remain in force and that no additional modifications are needed for GWAS.**