

## HEI POLICY ON THE PROVISION OF ACCESS TO DATA UNDERLYING HEI- FUNDED STUDIES

HEI has a longstanding commitment to producing science of the highest integrity, quality, and transparency that is built on a foundation of rigorous research and statistical design. HEI-funded research is subject to continuous oversight, data quality assurance audits, and more in accordance with HEI's Quality Management Program. The HEI Research Reports undergo intensive independent peer review, and *all* results are published. This *Data Access Policy* is designed to ensure access to underlying data for all HEI-funded studies.

The provision of access to data underlying HEI-funded studies is an important element of ensuring credibility, especially when the studies are used in controversial public policy debates. The open and free exchange of data is also an essential part of the scientific process. Therefore, *it is the policy of the Health Effects Institute to ensure access is provided expeditiously to data for studies that it has funded and to provide that data in a manner that facilitates review and verification of the work while protecting confidentiality and self-determination of any participants or communities involved in the study and respecting the intellectual interests of contributors to the original work.*

This policy applies to all research funded by HEI; it is consistent with The Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards (Uniform Guidance) at 2 CFR Part 200 of the Office of Management and Budget (OMB), which requires access under the federal Freedom of Information Act (FOIA) to data from federally supported research that was used in developing a federal agency action that has the force and effect of law.

In responding to FOIA requests through the U.S. EPA or other federal agency for HEI data that are subject to 2 CFR Part 200, HEI will follow the principles established in the amendments.

In responding to non-FOIA, direct requests to HEI for data, HEI will generally follow the principles described below, which are designed to be consistent with the requirements outlined in the 2 CFR Part 200 and the February 23, 2013 White House Office of Science and Technology Policy (OSTP) memorandum entitled "Increasing Access to the Results of Federally Funded Research", although specific cases might require other arrangements for providing access.

1. *Scientific Research Data* HEI uses a definition of scientific research data consistent with the one used in EPA's Policy for Increasing Access to Results of EPA-Funded Extramural Research ([link](#)) and the OSTP memorandum. The research data will vary from study to study, but generally will consist of the recorded factual material commonly accepted in the scientific community as necessary to replicate and verify the original research findings. It will include digital records including analytical summary and computer codes, where appropriate, but will not include any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. The "recorded" material excludes physical objects (e.g., laboratory samples). Research data also excludes (a) trade secrets, commercial information, materials necessary to be held confidential by a researcher until published, or similar information that is protected under law; and (b) personal and medical information and similar information that is personally identifiable, and the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.
2. *Scientific Data Management Plan and Provision of Access to Scientific Research Data* HEI will expect each Principal Investigator (PI) it funds to provide, at the outset of the study:
  - (1) a plan for organizing, protecting, archiving, and making all data, data descriptions, analytical summary, metadata, and computer codes described above available to HEI during or upon completion of the study to enable HEI's rigorous review process to ensure high quality of the results and to allow for independent quality assurance/quality control of the data, and
  - (2) a plan for making research data available to other investigators following publication of the results as set forth under this policy – preferably using public repositories that are made available under an open license– to allow for replication and reproduction of the results, as well as further exploration of the data by other investigators (see also *Timing* below).

In cases where all of the data used are from publicly available data sets and the analytic data set can readily and expeditiously be recreated, the PI might as an alternative provide detailed descriptions of how to access and use these public data sets to recreate the analytic data set in lieu of providing the full analytic data set.

3. *Third Party Data* In cases where the PI uses or purchases data for HEI sponsored research collected by a third party – whether public or private – and the PI is contractually bound with the third party to keep the data confidential, the PI will provide information on the process the third party has in place for access to the data and will direct the requestor to the third party to seek access to the data.

4. *Timing* HEI requires the PI to provide access to data as expeditiously as possible after the completion and publication of the HEI Research Report (or Reports) resulting from the study, taking into consideration the legitimate intellectual interests of the PI to have the opportunity to benefit from his or her intellectual endeavors and to publish subsequent analyses from the data set (including additional analyses funded by HEI). In some cases (e.g., for studies of particularly high regulatory importance being used to inform decisions over a short timeframe), the PI and HEI might need to work together to balance the PI interests against the need for interested parties to obtain access in a timely manner.
5. *Length of Data Retention* HEI-funded PIs are required to retain all data generated in the course of HEI-funded research for at least 5 years from the date of publication of the research by HEI, or a longer period if required by a funding agency or third-party data provider or as directed by HEI. HEI retains the right to access the data at any time during this period. If the PI has kept the data beyond this time, HEI will continue to have the rights to access to the data. At any point, and at least 90 days prior to any alteration or destruction or other disposal of the data, the PI will notify HEI so as to enable the Institute to request such data under this provision.
6. *Responsibility and Reimbursement for Costs* HEI encourages the PI to be the primary sharer of the data through public data repositories and to plan for effort and cost of data sharing at the outset of the study. This will reduce an undue burden on the PI in terms of providing data access to individual requesters after the study has been completed. To facilitate data access for all future and current studies in which HEI and the PI expect that the results have a high likelihood of being used in supporting a regulatory decision, HEI will consider requests from the PI for a reasonable budget of data archiving funds to be provided as part of the project budget.
7. *Confidentiality* In those cases when data have not been made available through public data repositories, any requester of research data will be expected to obtain any approvals and enter into any required data use agreements necessary to permit the requester access to such data. The requestor will be fully responsible for adhering to all such approvals from the appropriate agencies (e.g., the National Center for Health Statistics) or other third-party data providers. HEI will not knowingly itself provide, or require a PI to provide, information that can be used to identify a specific individual without the requester having already obtained all such necessary approvals.
8. *Responsibility of the Data Requester* In addition to the payment of reasonable costs and the obtaining of any necessary confidentiality approvals, HEI will ask the data requester, as would be normal courtesy in the scientific community, to inform both the PI and HEI promptly of any findings emerging from their analysis, to provide the PI an opportunity to respond to those findings prior to publication, to provide copies to both the PI and HEI of any papers submitted for publication from the data, and to cite both HEI and the PI in any publication, noting explicitly that the views expressed are those of the new analyst and not those of the PI, HEI, or HEI's sponsors.
9. *Data Availability Statement* HEI reports will include a data availability statement detailing the main project assets (research data, code, and results files) and how to access them. If certain data are not being made available, the Data Availability Statement will describe those data, the methods used by the investigators to access them, and the reason that they are not being made available.
10. *Data Repositories* PIs are encouraged to make their Scientific Research Data available through public repositories. The appropriate repository will depend on the specific scientific field of the study and may be selected in consultation with HEI. Wherever possible, widely used public repositories with an open license (e.g., Creative Commons) will be preferred.

The provision of data will not be simple to accomplish and will at times raise concerns and controversy from one or more parties. HEI will require access to data in a manner that to the maximum extent practical allows for replication of the research and fosters an atmosphere of collegiality and mutual respect among all parties, with the aim of obtaining from the sharing of data the maximum benefit for science and for the quality of the public policy decision-making process.