Policies for Use of Data from the Project Viva Epigenomic Data Repository

Requests for Data

These policies apply to use of **de-identified**, **epigenomic** data from Project Viva. Data may be individual-level or tabulated; individual-level data will be coded with Project Viva study ID. Any proposals requiring PHI or other identifiers (including dates) or genetic information will be governed by separate policies (see Project Viva Policies for Grant Applications, Analyses, Ancillary Studies and Publications/Authorship). An analysis plan must accompany all requests for use of Project Viva epigenomic data. To protect our participants and the quality and integrity of Project Viva data, Project Viva co-investigators will review all requests.

The Project Viva Project Manager will keep a database of all requests and determinations. Upon receiving notice of approval of a request for Project Viva epigenomic data, the Project Viva Project Manager will contact the proposing investigator to request the following:

1. Acknowledgement of and agreement to abide by Project Viva's Policies for Use of Data from the Project Viva Epigenomic Data Repository (this document).
2. A copy of IRB or equivalent approval of the proposed project by the investigator's home institution (if required by that institution), or a statement of confirmation that approval is not required.
3. A copy of the investigator's human subjects training certification, if required by his/her home institution.

After all of the above requirements are met, the investigator may request a dataset or summary statistics from the Lead Research Analyst. The Project Manager should be copied on this request. In situations requiring the expertise of another Project Viva-approved analyst, the Lead Research Analyst will provide a dataset to that analyst. Upon requesting and receiving Project Viva data (data sets or results) the proposing investigator agrees to follow all Project Viva policies as outlined in this document. He/she may use the data only for purpose originally requested. Approval must be granted by Project Viva investigators and the Harvard Pilgrim Health Care Institutional Review Board, as required, for additional use of the data.
This policy provides general guidance. Each proposal will be considered individually by Project Viva's operational leadership; specific requirements may differ from what is listed above.

Use of Data

Investigators may view and analyze only datasets or summary results that are provided to them by the Project Viva Lead Research Analyst or an approved Project Viva-affiliated analyst. Sharing a dataset or analyzing a dataset sent by another investigator, a lab, or any other individual or entity is against Project Viva's policies. In addition, investigators should not use old datasets for newly proposed analyses. The investigator will not attempt to re-identify observations included in an individual-level dataset, and will not merge the dataset with any other dataset not provided by a Project Viva analyst or approved as part of the analytic plan.

Viva's Lead Research Analyst does not have access to electronic locator files, medical record text files, or other operational folders that may contain any identifiable information or access to the linking code. Given these access limitations, she is not able to link data from the data repository to individual participants. Viva's Project Manager, Data Manager and other staff member may not disclose identifiable information or the linking code to the Lead Research Analyst or an outside investigator.

Project Viva's Lead Research Analyst is responsible for creating analytic datasets from the data repository and sending securely to the external investigator. She may also assist with analyses and be an author on papers if appropriate.

The investigator shall destroy Project Viva datasets created and distributed by the Lead Research Analyst within one year after resulting manuscripts are accepted for publication; this policy does not apply to summary results, which the investigator may retain as needed. The one-year timeframe allows the investigator to respond to any changes or re-run analyses based on the initial manuscript review.

To further safeguard participant privacy and confidentiality, the Lead Research Analyst will:

- Include only necessary variables in data sets.
- Email data sets using encryption.
- Email the data set with the below language.

  "The recipient has read Project Viva's Policies and agrees to abide by them. The recipient agrees to use or disclose the data only for the purpose requested, and for no other purpose. The recipient agrees to use appropriate safeguards to prevent any use or disclosure of the data. The recipient agrees to destroy any dataset provided by the Lead Research Analyst within the time frame outlined in these policies. The recipient will not attempt to re-identify the data or merge the data with non-approved datasets. The recipient will report to Project Viva's Project Manager any violation of this agreement or Viva Policies."

Manuscript Review

1. All manuscripts must be sent to the Project Viva Principal Investigator, Emily Oken, and Lead Research Analyst for review. Manuscripts may also require review by the Chair of the Department of Population Medicine (DPM) or the Director of Research at Harvard Vanguard Medical Associates' (HVMA) Office of Clinical Research; Viva's Lead Research Analyst will coordinate sending out the manuscript for these reviews if needed. Authors should allow 7 working days for DPM and HVMA reviews before submitting any manuscripts for publication; no response after this time can be considered approval. To promote collaboration within the group and shared knowledge of ongoing work, we will circulate the abstract of the submitted paper to the faculty (and staff) of the DPM's Obesity Prevention Program, with the disclaimer, "Draft—please do not cite or circulate". If for some reason

the lead author does not wish to have the abstract circulated to the wider group, s/he should inform the Lead Research Analyst.

2. All manuscripts involving individual-level Project Viva data must have a technical review to ensure that the data are presented accurately. In most cases this will be done by the Project Viva Lead Research Analyst. The technical reviewer will:
    a. Check for consistency and plausibility of numbers
    b. Double check tabulated numbers with primary SAS or R output
    c. Ensure that numbers add correctly etc.

Project Viva Staff Contact Information

Name	Title	Email	Phone
Sheryl Rifas-Shiman	Lead Research Analyst	Sheryl_Rifas@hphc.org	617-509-9824
Karen Ruderman	Project Manager	Karen_Ruderman@hphc.org	617-509-4925