



# **FAMILY HEALTH INDIA DATA PRIVACY POLICY**

**Family Health India  
H-5, Ground Floor Green Park Extension  
New Delhi-110016  
Phone: +91-11-26182107, 40487705**

**Finalized on: March 03, , 2021**

**Effective from: March 03, 2021**

**Table of Contents**

1 Introduction ..... 3

2 Access retrieval disclosure and use of traffic and message data..... 5

3 2.1 Principles of confidentiality and restricted purpose..... 5

3.2 Retention period and deletion of data ..... 6

3.3 Protection of retrieved data ..... 6

4 Personal Data ..... 8

3.1 Principles of good practice for processing or collecting Personal Data..... 8

3.2 Rights of Individuals ..... 8

3.3 Data collected by FH India for its own purpose..... 8

3.4 Data collected by customers of FH India ..... 9

4 Processing, storing and replication of traffic and message data ..... 10

4.1 Rules relating to processing and storing of Message Data ..... 10

4.2 Rules relating to replication of Message Data ..... 10

## **INTRODUCTION AND ACKNOWLEDGEMENTS**

Data are central to our work – whether it is the rigorous approach to collecting and synthesizing study data, the systematic monitoring and evaluation of FH India programs and services, or the application of data to inform policies, programs and services. Making more accessible the data collected through our scientific, technical, and programmatic efforts has the potential to inform future research agendas and the design and implementation of future programs and services.

Over the past decade there has been increasing international interest in making data, and datasets, more accessible. This move to open data requires organizations to apply the same rigor in sharing the data and datasets externally as in collecting, synthesizing and using the data in publications and in routine reports.

FH India’s response to open data, more specifically the creation of an internal working group and the shepherding of this document through initial drafts to a version that incorporates comments from around the organization.

It is important to note that these guidelines have been developed during a fluid time – a time in which the open data policies and requirements of Indian government and other funders are still evolving. As such these guidelines are based on the collective understanding of the working group at the time these were developed; we have deliberately kept them broad to allow flexibility to address individual requirements of funders and contracts. These guidelines will be reviewed periodically and updated as necessary.



## **WHAT DO WE MEAN BY DATA?**

Most funders define data to include any information collected during the period of performance of a project, qualitative or quantitative, that is in a structured format. This includes the data collected for research, as well as for indicator reporting, evaluations, assessments, and other monitoring and evaluation (M&E). Cost data, collected for the purposes of assessing or evaluating an intervention, should be included. However, open data policies generally do not include project financial information (e.g., FH India's accounting data) or data related to administration, such as equipment and supply.

Most open data policies specifically apply to data collected by the project. If the project conducts secondary analysis of existing data, the project is not responsible for submitting that data.



## CHECKLIST FOR ADHERENCE TO OPEN DATA POLICIES

Now that you know that open data applies to your project, here are the basic steps that you should undertake to ensure compliance. Additional guidance for each step can be found in this document as indicated in the column titled “relevant section.”

STEPS	RELEVANT SECTION	✓
Pre-award: Incorporating open data into proposals		
1.	Review funder policies and determine whether open data should be addressed in the technical proposal.	
2.	Budget for open data policy compliance beginning at the proposal stage.	
Post-award: Getting started with open data		
3.	Review funder policies and award-specific language with CMS at project kick-off.	
4.	Decide who on your project is responsible for implementing open data.	
5.	Discuss expectations with your funder contact.	
6.	Discuss open data with your local government partner. Some countries may discourage or prohibit sharing of data about their citizens or programs.	
7.	Discuss open data with implementing Partners.	
8.	Develop work plans, M&E plans, performance monitoring plans, and/or research protocols and informed consent forms with open data in mind.	
9.	Include costs associated with open data in budgets.	
10.	Develop an open data management plan (ODMP) that outlines the data to be collected, when and how it will be submitted, what documentation will be developed to accompany it, and which repository it will be submitted to. Include any requests for exemption for datasets that will not be submitted. 1. Submit to funder for review and approval.	
Post-award: Sharing data and submitting datasets		
12.	Ensure that participants consent to make data public.	
13.	De-identify data to ensure that participants remain anonymous.	
14.	Ensure/improve data quality through cleaning, review, and/or appropriate documentation.	
15.	Develop documentation to be submitted alongside your data.	



## 7.1 PRE-AWARDS: INCORPORATING OPEN DATA INTO PROPOSALS

Open data should be a consideration at the pre-award stage when proposals are being developed.

1. As a first step, the proposal lead should review the solicitation for any mention of open data. If there is no mention in the solicitation, also check the funder website for policies and guidance. If in doubt, inquire directly with the funder.
2. Begin planning compliance with open data policies as the project is designed in the proposal phase. Consider making note of plans to comply with the funder's open data policy, including any relevant details specific to how and where data will be shared, in the proposal text.
3. Some funders may require submission of an ODMP at the proposal stage.

## 7.2 HOW SHOULD WE BUDGET TO ENSURE COMPLIANCE WITH OPEN DATA POLICIES?

When developing a proposal (pre-award) and when refining and/or developing work plans and budgets (post-award), compliance with open data policies should be considered in the budget. This may mean new costs associated with preparing datasets and documentation for submission, as well as new and/or increased costs for ensuring data quality.

- i. Work with your finance/budget lead to include resources to cover anticipated costs. Any costs related to making your data publicly accessible should be allowable under the award.
- ii. Costs to consider include:
  - a. Staff time to develop an ODMP
  - b. Staff time for data entry, cleaning, and management, as well as de-identification of personal information
  - c. Staff time for development of appropriate data documentation for submission to a repository
  - d. Staff time and travel for data quality improvement activities, including training workshops, e-learning courses, and implementation of data quality and verification audits

Budget not only for project staff, but for support from other FH India teams and/or consultants, if assistance with data management, analysis, or quality will be needed.

- iii. There may also be costs associated with submission to privately managed data repositories, if you choose not to use, or do not have access to, a funder-managed repository. Confirm with the funder if those fees are allowable.

## 7.3 POST-AWARD: GETTING STARTED WITH OPEN DATA

Compliance with open data policies begins at the start of a newly funded project or study. This is true even when sharing of data will not take place until the final months of a multiyear project.



In the first few months of all new projects, FH India should engage in the following steps to ensure compliance with open data policies.

1. At the award kick-off meeting, key project staff should discuss open data requirements with FH India Contract Management (CMS) staff. CMS can review the contract language specific to the award and discuss the funder's policy.
2. The M&E Specialist will submit the dataset(s) to the appropriate repository. Note: The project director is ultimately responsible for ensuring compliance with open data policies.
3. Discuss expectations with your funder contact regarding making data from your project public.
4. For each dataset in the project, discuss who owns the data. Discuss with the local partner and funder contact at the outset of the project.
5. Open data policies "flow down" to Implementing partner. If you are hiring Implementing partner, discuss open data with them during the negotiation and contracting processes to ensure that they understand the need to adhere to open data policies. Project staff should work with CMS to ensure that subawards include open data language.
6. As initial project plans are developed and refined with the funder, ensure that work plans, M&E plans, performance monitoring plans, study protocols, and budgets are developed with open data compliance in mind.
7. Ensure that informed consent forms allow for open data. Suggested language is included in the informed consent form templates (See Annexure 1).
8. Include costs associated with open data compliance in budgets.
9. High-quality data are important to the success of open data and to Family Health India reputation. Plan for ways to ensure and improve data quality, starting from the beginning of the project. This may include getting technical assistance on the development and review of data collection forms and data entry and management systems; using electronic data capture; conducting data quality trainings with staff and partners; planning for regular data quality audits; and/or ensuring that you have budgeted adequate staff time for data cleaning.



## 7.4 POST-AWARD: DEVELOPING AN OPEN DATA MANAGEMENT PLAN

FH India recommends that all projects develop an ODMP at the beginning of each project, alongside development of the work plan, M&E plan, performance monitoring plans, and/or protocols. For some projects, development of an ODMP may start at the proposal phase. It may also be included as part of the work plan, M&E plan, or protocol.

As with a work plan or an M&E plan, the ODMP should be reviewed and approved by the funder. **This may or may not be a funder requirement, but FH India suggests it as a best practice to ensure that there is agreement between FH India and the funder on how the policy will be implemented by the project and to manage expectations.** The ODMP will be a living document, but changes should be discussed with and approved by the funder.

The ODMP will outline how the project will adhere to the funder's open data policy. At minimum, the ODMP should:

1. Describe all data to be collected, including what the datasets will cover, when and how they will be collected, plans for analysis, and how Family Health India will maintain and archive them.
2. Describe how informed consent will be collected and whether personal data will be collected and, if so, how the data will be de-identified (see Section 7.8 on data de-identification).
3. Include plans for when, how, and where the datasets will be shared/submitted, along with any documentation that will accompany the dataset.
4. If the project plans to submit the data to a repository other than the funder-managed repository, note which repository and why, and explain how that repository meets or exceeds the funder's requirements.

## 7.5 WHAT DATA WILL YOU SUBMIT?

The project manager(s) and M&E team should meet to discuss what constitutes data for the project.

1. List all datasets for the project. Include any that contain information you might put in a report or that might be the basis for a publication. Review project indicators. Include qualitative and quantitative data.
2. Agree on the level of data manipulation prior to repository submission. Options are to submit the raw dataset, a cleaned dataset, or an analysis-level dataset (with additional variables already calculated). For quantitative data, it is recommended that a cleaned dataset be submitted, but there may be instances in which another option is preferred. For most data repositories, there is also the option to submit multiple versions.
  - Check with your funder to see if there is a requirement for raw, clean, or analysis-level datasets.
3. Plan what documentation should accompany the datasets and who is responsible for preparing it.





4. If you have concerns about submitting any dataset, discuss this with your funder contact. It is recommended that you list the dataset, along with the concern(s) and/or justifications for not sharing the dataset, in the ODMP and request approval in writing from the funder.

When it comes to qualitative and mixed-methods data, deciding what to submit may be especially challenging. Use the ODMP as an opportunity to discuss options and come to agreement with the funder. At FH India, we are continuing discussion about what types and formats of qualitative data should and should not be submitted to open data repositories. Policies regarding qualitative data submission vary by funder.

## 7.6 WHEN AND HOW FREQUENTLY WILL YOU SUBMIT IT?

Different funders have different requirements for timing of dataset submission to a repository. For instance, policy states that a dataset must be submitted within 30 days of producing an “intellectual work” (e.g., a publication or report).

If the dataset is for indicators that are included in a quarterly report, you may be expected to submit an updated dataset quarterly. If multiple rounds of data collection are planned for a study or evaluation, such as baseline, midline, and end-line, you may be expected to submit data after each round.

Some funders may allow for data to be embargoed, or dataset submission to be delayed, beyond the required timelines if publications or patent submissions are planned. This should be agreed upon with the funder in writing in advance.

1. Learn about your funder’s requirements for dataset submission timelines and embargoes.
2. Document decisions about submission timelines, including planned requests for embargoes or delays, in your ODMP.

## 7.7 WHERE WILL THE DATA BE SUBMITTED?

Many of our funders, have their own open data repositories. In many cases, submitting to the donor-managed repository will be the best option. However, there may be situations in which submitting to a different repository has benefits. Some funders require data sharing, but do not have their own repository.

1. Check with your funder regarding repository requirements. Do they have their own data repository? Are you required to use their repository? What other repositories do they recommend or have been used in the past? Do they have requirements for how a repository must share the data?
2. Consider the pros and cons of using the funder-managed repository. Are there reasons to seek out a different repository? Funder-managed repositories are generally free to use but may have limited options for reviewing who has used your project dataset and why. The additional reviews of that dataset provided by the funder may be a benefit of using their repositories.
3. If interested, consider options for other repositories. Family Health India has developed a side-by-side comparison sheet of selected data repositories. This is not a comprehensive list but is based on those that staff have used previously or are familiar with.
4. If you use a repository other than the funder-managed one, you may still need to submit documentation to the donor-managed repository so that a link can be established from the funder to your dataset.



## 7.8 POST-AWARD: SHARING DATA AND SUBMITTING DATASETS

As you prepare to submit dataset(s) to an open data repository, there are a number of things to consider.

## 7.9 DO YOU HAVE PARTICIPANT CONSENT TO MAKE DATA PUBLIC?

In the past, informed consent language often stated that data would not be shared beyond the research team. Submission of datasets, even with Personally Identifiable Information (PII) removed, may be problematic in the context of the informed consent given by participants. Even if data are de-identified, there are risks of identification; participants have the right to know if their de-identified data will be submitted to an open data repository. You may need to obtain further consent from participants.

- Review informed consent language with the Indian Council of Medical Research (ICMR) or local Ethics Committee staff. If the original informed consent did not allow for data sharing:
  1. Discuss the situation and possible options with the ethics committee(s) that approved the work
  2. Discuss the situation and possible options with your funder contact

## 7.10 WHAT ABOUT DE-IDENTIFICATION OF DATA?

Before submitting a dataset to an open data repository, it must be de-identified, or have all PII removed. This means removing or obscuring any information that could result in individuals or small groups of people being identified from the data you share.

The process of de-identification requires a clear understanding of the variables collected, cultural and political context, size and uniqueness of the population, availability of related datasets that could be linked to the data, and risks involved in accidental disclosure. It does not merely involve removing names from a dataset.

The most difficult Personally Identifiable Information (PII) is indirect, such as when a combination of variables can lead to identification of a participant. This is especially true if the number of participants is small and/or drawn from a particular population (e.g., the population consisted only of Ministry of Health staff).

Qualitative data are especially challenging to de-identify, as the sample sizes are generally small. Given the personal and narrative nature of most qualitative data, identifiable information is likely inextricably woven throughout. This makes redacting all identifiable or sensitive information from transcripts, interview notes, and other raw qualitative data sources extremely difficult. If coded qualitative data summaries are shared, these too should be reviewed carefully to ensure that identifiable information is removed or masked.

1. M&E, data, and/or research staff should meet to discuss the identifiable variables in the dataset, or those that could lead to identification. Develop and agree on a plan for removing or obscuring potentially identifiable data variables.
  - i. To mask data, consider grouping data within variables. A few examples of this are:



- Age: use age ranges such as 18-24 years, 25-34 years, or 35-44 years.
  - Dates: remove exact dates and default to the first of the month.
  - Geographic areas: remove district-level data, using only region.
- ii. For large or complex datasets, data masking software can also be deployed.
2. Implement the plan for de-identification and then have the M&E, data, and/or research staff review the dataset again. Has all Personally Identifiable Information (PII) been removed?
  3. If time and resources allow, ask for an additional review to ensure that PII has been removed. Some funders/repositories will provide an additional review before publishing.

### 7.11 WHAT ABOUT DATA QUALITY?

As an organization that prides itself on high-quality work, FH India would like to ensure that datasets made publicly available reflect positively on the organization. While maintaining a high standard of quality across FH India's many datasets will be challenging, FH India asks for project teams to make every effort to ensure that their data products meet a high standard.

As time and resources allow, there are several ways to improve data quality. This should begin as early in the project as possible.

It is advised that projects have completed datasets, along with the supporting data documentation, reviewed for quality before submission to an open data repository. Review should be provided by an experienced data analyst at FH India. This review should provide recommendations for improving the quality of the dataset and the usefulness of the data documentation, as well as comments on the quality issues to be noted in the data documentation.

If a dataset is not considered high quality (due to lack of time, resources, or capacity), or if the funder insists that you submit a partial or uncleaned dataset, indicate these quality concerns clearly and openly.

### 7.12 WHAT DATA FORMAT SHOULD YOU USE?

Check with your Donor and/or planned repository to see if there is a requirement for dataset format or program. Some funders/repositories require that a dataset be submitted in a nonproprietary format, such as comma separated values, as opposed to proprietary software such as Adobe, Microsoft, or SAS. You may also be able to submit your dataset in multiple formats to facilitate use and to prevent errors in conversion by secondary data users.

If your data are in a proprietary format, use a conversion program like Stat Transfer to convert them to comma separated values files.

### 7.13 WHAT DOCUMENTATION SHOULD BE SUBMITTED ALONGSIDE DATASETS?

Data made openly available will not be useful if there is not enough context provided to enable someone else to use it accurately and appropriately. Documentation submitted alongside the dataset should include a basic description of the study or project, including the data collection design and methodology, a codebook or data dictionary defining the variables, a description of the relationships among the files within a dataset (if more than one), and any other relevant



contextual information to help others accurately interpret the data. If the data will be updated over time (e.g., for a regular indicator reporting), make note of that in the documentation.

A data quality cover page is being developed for inclusion in the documentation for data collected and submitted to an open data repository by FH India. This page explains how FH India works to ensure high-quality data, allowing comments and concerns specific to the dataset to be noted.

Some project teams may also choose to submit an analysis plan to enable secondary users to more accurately replicate their analysis; this is not required. Some project teams, particularly for research studies, may feel that submitting an analysis plan would mean giving up FH India's intellectual property.

For research studies, appropriate documentation may have already been developed to support data collection, cleaning, management, and analysis. For programmatic data, however, this documentation may exist in M&E plans, performance monitoring plans, informal documents, and/or as disparate notes and will need to be formalized for submission.

For qualitative data, choosing the appropriate level of data and accompanying documentation to include may be more challenging. Review the funder policy and discuss options with your funder contact.

Data documentation should also provide clear guidance on how to cite the dataset, and whom to contact for additional information.

## USE OF OPEN DATA

### 7.14 HOW DO WE HANDLE OTHERS' USE OF DATA COLLECTED BY FH India?

Once datasets are available through an open data repository, it is likely that they will be used by others. While there is a risk of data being misused, in most cases, the data use will be positive because open data can create efficiencies, prevent duplication of research and data collection, support the translation of data into evidence-based programs and policies, and allow for study replication and the triangulation of data for answering new questions. These uses extend the positive impact of the data collected.

- In the documentation submitted with any dataset:
  - Provide a contact at FH India M&E Specialist (Mr. Rajesh Kumar Mishra Email: [RMishra@fhiindia.org](mailto:RMishra@fhiindia.org)) who can answer questions about the dataset. Emails received at this address will be reviewed by Knowledge Exchange staff who will identify the most appropriate person.
  - Request that any user of the data inform FH India of the data/study, purpose, and any publication resulting from its use.
- Project staff should keep track of all secondary uses of their datasets, to the extent possible, and report them to the funder as outcomes of the project.
- If FH India is contacted with questions from people who are using our open data, discuss the situation with project or departmental leadership. If relevant, it might also be useful



to discuss with the funder contact. Carefully consider the pros and cons of responding, including whether staff have the time and resources to respond.

### 7.15 WHAT ARE FH INDIA'S EXPECTATIONS FOR STAFF USE OF OTHERS' DATA?

FH India expects its staff to use data from publicly accessible datasets in an ethical manner with respect for context and anonymity, and with reference to those who collected the data.

- If you are using secondary data for research purposes, follow FH India's policies and procedures.
- Read all documentation for the dataset to understand the methods and population.
- Seek clarification from the dataset contact when you do not understand something.
- Respect anonymity and the confidentiality of the data.
  - Review the documentation to confirm that consent was provided to share the data publicly.
  - If you find a dataset with identifiers on it, report it to the owner of the data repository.
- Respect context; be careful when extrapolating results beyond their intended population or geography. When you do so, be clear about the limitations of your interpretation.
- Cite the data appropriately. Use the suggested reference/citation language provided, or in the absence of suggested language.



**ADD TYPE OF INSTRUMENT**  
**ADD NAME OF PROJECT**

Namaste. My name is ..... I am working with (name of organization). We are conducting research of the households ----- (name of project and geographies where the research will be conducted). We are interviewing many households about their knowledge on ----- (topic of interview). Your family was randomly selected to be interview. You have the right to participate in the survey or to refuse participation. If you choose to continue with the interview, you can choose to respond to any or all the questions we ask. Your responses will be used for research purposes only. We will not tell anybody what your individual responses were. Your name and your answers will not be disclosed to anyone. Kindly note that data gathered from your response will be de-identified (will not have any personal identifiable information) and might be shared on an open data repository.

The length of this interview is approximatively thirty minutes. You will not receive immediate benefits from participation in the survey, although the results of this research will be used for ----- (objective of the project).

Can we continue the interview?

If you feel you have been treated unfairly, or you have questions or concerns you may contact ..... (Name of Principal Investigator and mobile number)

**ORAL CONSENT OF RESEARCH SUBJECT**

I understand the procedures described above. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

\_\_\_\_\_  
Name of the Subject

Oral consent provided 1. Yes 2. No

(Circle the answer that applies)