



Haramaya University

**Kersa Demographic Surveillance and Health Research
Center (KDS-HRC)**

Data Sharing and Access Policy

Oct.2014

Harar, Ethiopia

Acronyms

CD-ROM	Compact Disk-Read Only Memory
ID Number	Identification Number
INDEPTH	International Network for the Demographic Evaluation of Populations and Their Health
IRB	Institutional Review Board
IT	Information Technology
KDS-HRC	Kersa Demographic Surveillance and Health Research Center

Table Contents

Title	Page
I. Introduction	1
II. Rationale for KDS-HRC Data Sharing Policy	2
III. Goals of KDS-HRC Data Sharing Policy	2
IV. Goals of KDS-HRC Data Sharing Policy	3
V. Benefits of the Policy	3
VI. Legal framework	4
VII. Data Sharing Methods	4
VIII. Levels of Access	4
IX. Data sharing terms of agreements	5
X. Data Sharing Proposal Review	7
XI. Updating of the Policy	7
XII. Ratification of the Policy	7
XIII. Fees	7
XIV. Handling of breach of the Policy agreement	8

I. Introduction

Data sharing is the practice of making data used for research available to other investigators or data users. Scientific research depends on the free flow of information and ideas. To ensure that future research can build on previous efforts and discoveries, the KDS-HRC developed a data sharing policy that goes into effect after the endorsement of the document. The underlying principle in developing this data access and sharing policy is that data are a public good and sharing data is both ethical and beneficial. The policy guides the final community research data to be made available to other investigators and sets out the specific procedures and access levels related to various categories of data covered under the policy.

Protecting the rights and privacy of human subjects should be the first priority of any research undertaking. Investigators, Institutional Review Boards, and research institutions have an obligation to protect participants' rights and confidentiality. However, data sharing is possible without compromising these efforts because identifiers can be removed from data. In addition, data sharing agreements can be used to restrict the transfer of data to others and to require that data be used only for research purposes.

An important issue associated with the sharing of all data derived from human subjects is the protection of research participants' identities. The rights and privacy of people who participate in any type of research should be protected at all times. Sensitive data raise special concerns about confidentiality and the protection of subjects' privacy because of a greater possibility of harmful social, economic, or legal consequences if released. They call for a higher level of security during collection, analysis, and storage and special consideration when preparing datasets for broader use. What constitutes "sensitive" data varies by context, population, and time. Illegal and sexual behaviors are almost always considered sensitive. Measures of alcohol use are less sensitive among adults than underage adolescents. Many diseases and medical conditions, such as psychiatric illness or HIV infection, could be considered sensitive information.

There are two basic tools to protect from disclosure of sensitive data and subjects' identities: Restricting information in the dataset, and restricting access to the data. Thus, data intended for broader use should be free of identifiers that would permit linkages to the research participants and free of content that would create unacceptably high risks of subject identification. In essence, shared datasets must be anonymised and should not contain data elements that will allow the direct identification of the entity (individual, household or locations) represented by a data record, such as names, identification numbers or geo-coordinates. It is not possible to completely eliminate the risk of indirect identity disclosure where a combination of data elements may be used in combination with (linked to) other information known to an intruder about the entity in question to identify the entity. Access to datasets should be assessed based on the disclosure risk posed by the dataset. Only datasets that pose a low risk of identity disclosure should be made publicly available. In most practices data access reviewers develop guidelines for the assessment of identity disclosure risk in data to be shared. Such guidelines should include recommendations for specific access restrictions to datasets depending on the measured disclosure risk.

Stripping a dataset of items that could identify individual participants is referred to by several different terms, such as data redaction, de-identification of data and anonymizing data. It is rarely sufficient to simply remove names, house addresses, identification numbers, geographic

information and other features or elements of dates such as those for birth, hospital admission and discharge, death and the like. Deductive or indirect disclosure of individual subjects may become more likely when there are unusual characteristics or the joint occurrence of several unusual variables. Samples drawn from small geographic areas, rare populations, and linked datasets can present particular challenges to the protection of subjects' identities.

II. Rationale for KDS-HRC Data Sharing Policy

There are many reasons to share data from KDS-HRC. Among these, it:

1. Reinforces open scientific inquiry,
2. Encourages diversity of analysis and opinion,
3. Promotes new research,
4. Makes possible the testing of new or alternative hypotheses and methods of analysis,
5. Supports studies on data collection methods and measurement,
6. Facilitates the education of new researchers,
7. Enables the exploration of topics not envisioned by the initial investigators, and
8. Permits the creation of new datasets when data from multiple sources are combined.

III. Goals of KDS-HRC Data Sharing Policy

Data sharing promotes many goals of the KDS-HRC research endeavor. It particularly is used as a secondary data source for investigators who cannot readily obtain or access such longitudinal data by themselves. However, its ultimate goal is to promote the wide use of its data by the scientific community and eventually make more meaningful contribution towards promoting societal wellbeing.

In KDS-HRC's view data should be made as widely and freely available as possible while safeguarding the privacy of participants and the study community at large, and protecting confidential and proprietary data. In essence, KDS-HRC's data sharing and access policy aims to

1. Provide a platform for proactive and open access to data generated through a continuous follow up of demographic and health related events.
2. Reduce practical and financial limitations in obtaining and processing large data sets.
3. Maximize the number and types of analyses performed on scarce resources.
4. Minimize unnecessary duplication of effort.
5. Create a standardized methodological and informational infrastructure.
6. Provide oversight to ensure adequate protection of participant, community and proprietary rights.
7. Give clear understanding on the basic principles of sharing data.
8. Ensure beneficence of the study community.

IV. Scope of the Data Sharing Policy

The KDS-HRC data sharing and accessibility policy applies to the sharing of all community data and information collected, generated and archived by the center.

V. Benefits of the Policy

In principle this policy is expected to strengthen the Centre's capacity to effectively manage, curate, analyse community data and publish research outputs as well as to foster balanced collaborations and partnerships with the scientific community. More on that, the following are the core benefits of KDS-HRC's data access and sharing policy.

1. **Maximizing use:** Ready access to KDS-HRC owned data will enable more extensive use of a valuable public and institutional resource for the benefit of the data user and the community at large.
2. **Avoiding duplication:** By sharing data the need for separate bodies to collect the same data will be avoided resulting in significant cost savings in data collection.
3. **Maximized integration:** By adopting common standards for the collection and transfer of data, integration of individual data sets may be feasible.
4. **Better decision-making:** Data and information facilitates making important decisions without incurring repetitive costs. Thus, ready access to existing valuable data that are principally owned by KDS-HRC to the decision makers is essential for many decision making tasks such as promoting health education programs, protecting the environment, development of health planning, managing assets and resources, improving living conditions, and controlling epidemics etc.
5. **Equity of access:** A more open data transfer policy ensures better access to all bona fide users.
6. **Beneficence:** Ensures maximum benefit to the study community from the data and information generated from it.

VI. Legal framework

Data remains principally the property of Haramaya University KDS-HRC, which originally collected or created them and resides in its IT enabled facility for sharing and providing access. Access to data under this policy should not be in violation of any Acts and rules of the Government of Ethiopia as well as the Oromiya National Regional State and Harari Region where the study area is located in force. Legal framework of this policy will be aligned with various Acts and rules covering the data.

VII. Data Sharing Methods

KDS-HRC recognizes that different approaches to data sharing will be required in different situations and considers that it is most appropriate for researchers to determine their own strategies for data sharing and outline these within their data sharing application.

The KDS-HRC's managed access data sharing can be accomplished through:

1. Publishing (articles in scientific publications),
2. Responding directly to data requests (delivering CD-ROM or hard copies containing data)
3. Data Enclave—In this method data can be shared under the most controlled and secured environment in which eligible researchers can perform analyses using data resources. If, for example, there is any risk of subject identification, the center may ask that users submit requests for specific analyses or come to the site to run analyses under supervision. No electronic or hard copies of data can leave the center unless approved by authorized KDS-HRC staff. Fee will be charged for work at the data enclave that includes space, equipment, staff time for support, supervision and analysis, as well as the creation and maintenance of data files required by the researcher. In addition, researchers must work under the supervision of KDS-HRC staff during normal working hours.
4. Mixed Mode—more than one version of a dataset, each providing a different level of access can be used.

VIII. Levels of Access

A prospective data user can access the data up on submission of application together with statement of purpose for which the data will be used and an agreement to the conditions of data use as detailed in this policy document is signed.

The following access levels are applicable for KDS-HRC data sharing:

1. **Open Access:** The data will be made available without any restriction; for instance, providing sampling frame.
2. **Restricted Access:** Any data request other than sampling frame and highly sensitive data but which includes request of non sensitive data that passed maturation period¹ (two years) for detailed analysis.
3. **Closed Access:** This applies to highly sensitive or individually identifiable data that passed maturation period (two years) for detailed analysis. Such data are normally available to prospective users only through controlled-on-site access in a data enclave.

¹ Maturation period is the time that the Center uses for data cleaning and preparation for rigorous analysis.

Many diseases and medical conditions, such as psychiatric illness or HIV infection, sexual behaviors could be considered sensitive information. Access to geo-coordinates are also considered in this category.

IX. Data sharing terms of agreements

Responsibilities and Rights of KDS-HRC

1. Data released for sharing by the research centre should be validated and verified in line with accepted best practice and be of high quality. In so doing, the centre should properly document the longitudinal data to ensure that others can use the dataset and to prevent misuse, misinterpretation, or confusion.
2. The research centre shall give data important for sampling frame to researchers interested to conduct a research in the site and for teaching/learning purposes as per the agreement there of . Based on the initial signed agreement data can also be shared with the network of other DSS sites in Ethiopia or elsewhere and with INDEPTH networks as well as with all appropriately qualified researchers from academia, charity organizations and private companies.
3. The research team member or any staff facilitating data transfer shall not require co-authorship as a condition for giving data. But if their technical/professional contribution is appreciated by the researchers, they can be included in the list of authors in the form of joint authorship.
4. Any managed access procedures of the centre should be consistent and transparent. The data access committee who review or assess applications to access data should have appropriate expertise and must be independent of the project to be reviewed.
5. When making data set available, it may place limits or set conditions or establish a graded access procedure as necessary depending on the type and sensitivity of data required.
6. Shared datasets by the centre must be anonymised and should not contain data elements that will allow the direct identification of the entity (individual name and ID). Geo coordinates will be accessed in a data enclave.
7. Timescales for data sharing will be influenced by the nature of the data and since KDS-HRC collects its longitudinal data over several discrete time periods or waves, it is reasonable to expect that the data would be released in waves as data become available or main findings from waves of the data are analysed. However, it is expected that the research centre's timely release would generally be no later than the release through publication of the main findings or not more than two years of the generation of the data.
8. The research centre will not be responsible for any claims that the application of information generated from the data lead to wrong conclusions/decisions by the data user or other third party. Neither the collaborators to KDS-HRC bear this responsibility.
9. The research centre requests fees to recover costs of making data available for data users and for the community beneficence. However, students and staff of Haramaya University may be exempted from this cost provided that they have not received any fund for this activity.
10. The center requires the data user to submit a written manuscript and policy briefs ready for publication according to the submitted objective and title in not more than three months. The sample for policy brief will be given from KDS-HRC coordination office.

11. The center would make necessary assessment to the manuscript and give its comment to the data user in not more than one and half months.
12. Failure of the data user to submit a manuscript in three months would trigger KDS-HRC to withdraw its agreement with the data user and communicates to concerned bodies for necessary disciplinary actions.
13. The data remain the property of the KDS-HRC which reserves the right to request the return of the dataset should any of the conditions in this policy document be violated.

Responsibilities and Rights of data user

1. The data user should protect the privacy of subjects and the confidentiality of the data as well as observe the ethical and legal obligations pertaining to the data. To this effect, a signed confidentiality agreement form should be completed by him/her to access the data set. In addition, the data user should secure the data using appropriate computer technology.
2. The recipient of the data should be refrained from the act of manipulation of data for the purposes of identifying subjects.
3. The recipient shall not transfer the data to other users in any form with or without imbursement; that the data are only to be used for IRB reviewed research or for teaching/learning purposes. It should be noted that the identified data can be disclosed or used for a purpose other than that for which it was supplied only if the KDS-HRC coordinator has consented with a written and signed agreement.
4. Data users must submit a detailed description of their projects to the research center to gain access to restricted and closed data.
5. The proposal must include personal identification and institutional affiliation, a current resume, a detailed summary of the proposed research including a statement of why the KDS-HRC data are needed, the intended use of the data including a list of expected outputs, and a complete list of data requested including data type and volume, content, variables, years, and the like. The KDS-HRC staffs are expected to avail themselves for consultation on the data sharing proposal development.
6. All the data users who are accessing / using the KDS-HRC's data shall acknowledge the center in all forms of publications. Any books, articles, conference papers, theses, dissertations, reports or other publications employing data obtained from KDS-HRC should cite the source, in line with the citation requirement provided with the dataset.
7. The investigator or data user should send an electronic copy of all manuscripts with policy briefs/publications based on the requested data to KDS-HRC not later than three months for manuscripts and as published for publication from restricted and closed access data.
8. Data users should destroy or return the data after analyses are completed.

X. Data Sharing Proposal Review

Data access committee consisting of at least six KDS-HRC staff with appropriate expertise and who are independent of the project will review all the proposals. This review may address the following critical questions: Does the proposed activity constitute statistical research? If so, is there any risk that respondents will be identified inadvertently? To what level of access the data should be availed? etc. The committee should keep a record of data access requests which includes: data set(s) requested, date of application, minimum applicant information including name, contact e-mail/address, and institutional affiliation, relevant publications, whether the applicant's proposal was approved or not, date of approval/denial of access and reason(s) for denial of access.

XI. Updating of the Policy

This policy will be up dated every five years on regular basis or as deemed necessary.

XII. Ratification of the Policy

The policy will be effective after endorsement/ratification by the Research Directorate of Haramaya University.

XIII. Fees

The KDS-HRC will charge data recipient depending on the level of access:

1. For sampling frame 5% of the total research budget shall be charged to staff and students of Haramaya University. For external users the charge will be 15% of total research budget.
2. For restricted access 10% of the total research budget shall be charged to staff and students of Haramaya University. For external users the charge will be 20% of total research budget.
3. For closed access 20% of the total research budget shall be charged to staff and students of Haramaya University. For external users the charge will be 35% of total research budget.

XIV. Handling of breach of the Policy agreement

Violation of this policy agreement in general and failure to comply with any of the above requirements in particular will be handled by the agreement signed between the two parties and other appropriate laws governing such activities.

Made into effect by:

KDS-HRC, Haramaya University

Ratified by:

Research Director, Haramaya University
