Ethiopian Universities
Health and Demographic Surveillance System

Network Data Sharing and Release Policy

September 2013

Table of Contents

Abbreviations

AAMSP: Addis Ababa Mortality Surveillance Program

AM-DHDP: Arba Minch Zuria Demographic and Health Development Program

BRHP: Butajira Rural Health Program

CDC: Centers for Disease Control and Prevention.

DHS: Demographic and Health Surveys

DRC: Dabat Research Center

EPHA: Ethiopian Public Health Association

GGFRC: Gilgel Gibe Field Research Center

HDSS: Health and Demographic Surveillance Systems

INDEPTH: International Network for the Demographic Evaluation of Populations and Their Health in Developing Countries

Ishare: INDEPTH Access and Sharing Repository

KAD-HDP: Kilite Awlaelo Demographic and Health Surveillance Site

KDS-HRC: Kersa Demographic Surveillance and Health Research Center

MOA: Memorandum of Agreement

NFFS: National Family and Fertility Survey

NSC: Network Scientific Committee

SNNPR: Southern Nations and Nationalities and Peoples
            Region

SPSS: Statistical Package for the Social Sciences

VA: Verbal Autopsy

Acknowledgments

1. Introduction

The generation, dissemination and utilization of demographic and health related evidence through census, vital events registration, population based surveys and surveillance system, are fundamental for evidence based teaching, policy formulation and decision, program planning and practice. Countries generate these types of information through different combinations of the above methods, depending on the type and level of the required information and the resources available to them. Most developed countries conduct vital events registration, census and national population based surveys and surveillance systems on regular basis.[1,2] Unfortunately, such evidences are generally limited in developing countries mainly due to resource constraints. In Ethiopia, one of the major sources of information is the Population and Housing Census which is conducted roughly every ten years. So far, such censuses were conducted in 1984, 1994 and 2007. The population count during the last census was 73.8 million while growth rates over the three periods were 3.1% per annum in 1984, 2.9% in 1994, and 2.6% in 2007.[3] The other sources of information in the country included passive surveillance from health institutions,

the 1981 Demographic Survey, the 1990 National Family and Fertility Survey (NFFS), the 2000, 2005, and 2011 Demographic and Health Surveys (EDHS). The 1990 NFFS was the first nationally representative survey which yielded substantial information on fertility, family planning, contraceptive use, and related topics. In addition to the topics covered by the NFFS, the 2000, 2005, and 2011 EDHS surveys collected information on maternal and child health, nutrition and breastfeeding practices, and HIV and other sexually transmitted diseases. [4, 5]

Although these nationwide surveys were used as sources of information, they were conducted infrequently. Moreover, there has been no continuous systematically organized registration of vital events. In addition to their main purpose of integrating teaching, research and community services, some universities in Ethiopia have established Health and Demographic Surveillance Systems (HDSS) at different corners of the country in order to contribute to filling the information gap to some extent. Currently, six HDSS sites and one mortality surveillance program are networked to produce data which may reflect some pictures of the country related to health and demography. This network namely the 'Ethiopian

Universities Research Centers Network' was established in 2007 with the support obtained from the Ethiopian Public Health Association (EPHA) and The Centers for Disease Control and Prevention (CDC). When the network was established, the first common activity was investigating causes of death using the verbal autopsy method. Now, the main purpose of this collaborative endeavor is to support universities in generating longitudinal data that will be useful for strategic information for the country and university researchers and non-university communities in addition to VA activities. In addition, these HDSS sites help to build the research capacity of staff and students. For example, these sites are being used as sites where university students have hands-on training on the design and conduct of surveys and surveillance. This collaborative undertaking has two components, Verbal Autopsy and population updates through the registration of events like birth, death, marital status change, migration, pregnancy, and housing condition following the baseline census.

The networked research centers include the Butajira Rural Health Program (BRHP) and the Addis Ababa Mortality Surveillance Program (AAMSP) owned by Addis Ababa University; Dabat Research Center (DRC) owned by the University of Gondar; Gilgel Gibe Field Research Center (GGFRC) owned by Jimma University; Kersa Demographic Surveillance and Health Research Center (KDS-HRC) owned by Haramaya University; the Kilite Awlaelo Demographic and Health Surveillance Site (KAD-HDP) owned by Mekele University, and the Arba Minch Zuria Demographic and Health Development Program (AM-DHDP) owned by Arba Minch University.

The Butajira Rural Health Program (BRHP) was established in 1986 with the collaboration of the then Community Health Department, Medical Faculty of Addis Ababa University and the Epidemiology and Clinical Medicine, Umea University, Sweden as a PhD project. It is located in the Guraghe and Silti zones of the Southern Nations and Nationalities and Peoples Region (SNNPR). This surveillance site has one urban and nine rural kebeles that were selected randomly out of four urban and 82 rural Kebeles of the then Meskan and Mareko district.[6]

The Dabat Research Center (DRC) was established in 1996. Dabat town, the capital of Dabat District, is located approximately 821 km Northwest of Addis Ababa and 75 kms north of Gondar town. This site covers 10 randomly selected Kebeles (seven rural and three urban) out of the 32 Kebeles in the district.[7,8]

The Addis Ababa Mortality Surveillance Program (AAMSP) was initiated in February 2001 as part of the Master of Public Health study in the School of Public Health, the then Department of Community Health, Addis Ababa University. A prospective surveillance of burials was established at all cemeteries within the city limit with a primary objective of monitoring the impact of HIV/AIDS on mortality. Since July 2006, EPHA in collaboration with CDC, has funded and administered the program with the objective of monitoring the population level impact of HIV/AIDS on mortality and evaluation of antiretroviral treatment intervention.[9]

The Gilgel Gibe Field Research Center (GGFRC) was established in 2005 comprising 11 Kebeles (eight rural and three urban) bordering the Gilgel Gibe Hydroelectric Dam within about 10 km radius covering four Woredas (Sekoru, Tiro Afeta, Omo Nada and

Kersa). The center is located about 260 km southwest of Addis Ababa, and 55 km northeast of Jimma town.[10]

The Kersa Demographic Surveillance and Health Research Center (KDS-HRC) was established in 2007. Kersa Woreda is located in Eastern Hararge zone of the Oromia Regional State in eastern Ethiopia. The Woreda capital, Kersa, is located 44 km from Harar in the western direction. The KDS-HRC site covers 12 randomly selected Kebeles (two urban and ten rural) out of 38 Kebeles in the woreda.[11]

The Kilite Awlaelo Demographic and Health Development Program (KAD-HDP) was launched in 2009 and is comprised of 10 Kebeles (one urban and 9 rural) that were randomly selected from three Woredas in Tigray Regional State. Kilite Awlaelo is located approximately 802km from Addis Ababa.[12]

The Arba Minch Zuria Demographic and Health Development Program which was opened in 2009 comprises of 9 Kebeles (one small town and 8 rural Kebeles) randomly selected out of 29 Kebeles in the Woreda. Arba Minch Zuria Woreda is located in Southern

Nation, Nationalities and Peoples Region (SNNPR). Arba Minch town is located approximately 502 km south of Addis Ababa.[13] The locations of the different research sites are shown in Figure 1.
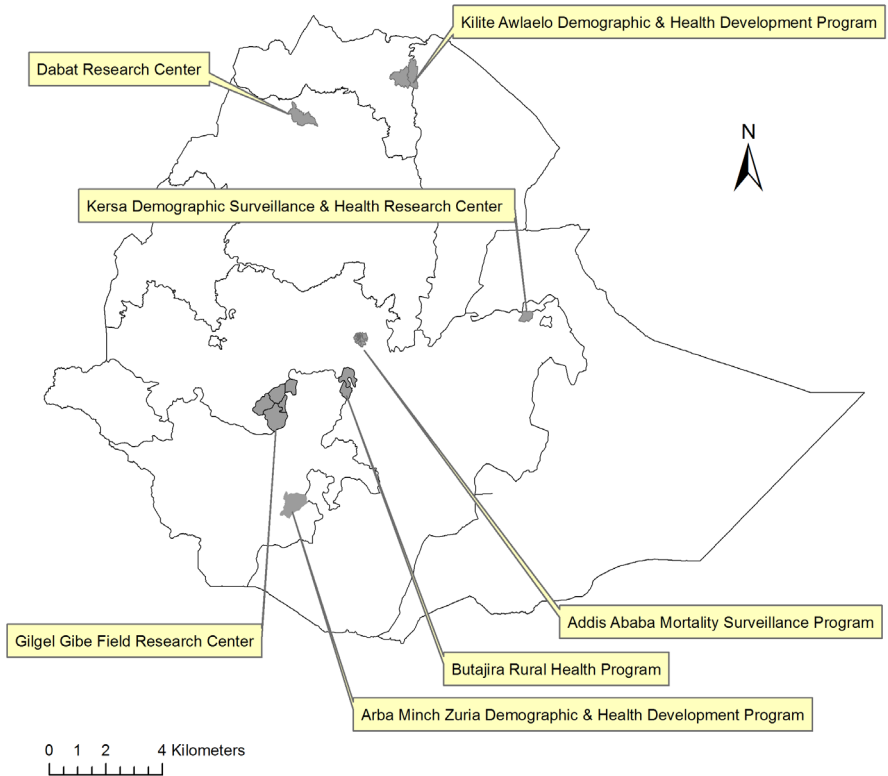

Figure 1: Map of Ethiopia showing the location of research sites

2. Rationale for the data sharing and release policy

While BRHP updates population every three months, the other HDSS sites do this biannually. The member research sites produce their own detailed reports or publications from the data they collect. The depth and frequency of such reports or publications is decided by the respective sites. In addition to site specific reports, it has also been agreed to produce network-level reports by combining the data from member sites. Though there is a general consensus on producing network level reports, the details, including what data each member site shall contribute is not yet worked out. In fact, when agreement is entered between EPHA and respective sites every year, part of the agreement is to send clean data to EPHA based on the agreement format. It is also wise to share and release data to other individuals/groups/institutions for efficient utilization of the data stored at the respective research sites and network level. Even though there is an interest to share and release data, there has been no clear guideline on how to implement it. That is why this policy is developed. This policy will also be applicable to sites which may join the network in the future. New sites

in the future can register as members of the network only if they satisfy the requirements of the network and the funding agencies

3. Objective of the policy

The objective of this data sharing policy is to guide the effective, efficient, and appropriate/ethical utilization of data generated by the networked sites in Ethiopia through the harmonization of national and collaborating partners' requirements and guidelines.

4. Types of data that can be shared and released

4.1 Levels of access to data

The network can consider the following levels of access

4.1.1 Open Access

Data will be freely available for every one who is interested to use. Getting permission from the network is not necessary for the utilization of such data which are identified to be open for every user. Such data is released through a public use data set

4.1.2 Restricted access

Prospective data users shall sign an agreement with the network before data is shared to them.

4.1.3 Closed Access

Data will not be shared with any one for different reasons like the sensitive/confidential nature of the data.

4.2 Minimum sets of data that can be shared to the network by the respective member research sites

• Data necessary for describing the socio-demographic characteristics of the study population e.g age & sex dis tribution of the study population

• Data necessary for calculating different fertility rates

• Data necessary for calculating different mortality rates

• Data necessary for calculating different migration rates

• Data necessary for calculating rates of marital change

• Data necessary for analyzing the causes of death

• Other data in times of national level requests like vaccine trials

•Any scientific report produced by each site shall be sent to the network for documentation

•Member sites shall develop data sharing and release

policy which is in line with the network level policy and the mission, vision and mandate   of the respective universities.

4.3 Types of data that the network can share to other data users (Restricted access)

- The network level scientific committee shall decide the type of data that will be shared
- The committee shall develop data sharing templates with defined common variables.
- Agreement shall be entered between the network and the prospective data users with regard to the type of data that will be shared

4.4 Types of data that can be released by the network

While releasing data is important, the network should develop strategies of maintaining the confidentiality of the information collected from the study units. The network shall also develop formats which will be used for releasing data. The necessary details of the data need to be issued to understand the data e.g data dictionary. The network scientific committee will continuously decide the type of data that can be released.

5. Implementation strategies of data sharing and release

5.1 Coordination of network data related activities

- Since the network is under the EPHA legal entity, EPHA will coordinate the network data related activities.
- EPHA shall work very closely with each site registered as a member
- EPHA shall establish a data warehouse, maintain and manage data and commit the required material and human resources at the EPHA Head Office.
- EPHA shall be responsible for the establishment of the Network Scientific Committee (NSC), facilitate and follow its functionality
- EPHA shall serve as secretary of the NSC
- Each site shall submit data to the network data warehouse with agreed sets of variables, procedures, tools andtimeline.
- Each site shall establish and maintain a site level data warehouse.
- Each site may develop its own data sharing and release

policy that shall be in line with the network policy.

- The network shall ensure the overall standardization of data collection tools, procedures and methods.
- NSC under the network shall ensure data quality to be shared.
- NSC under the network shall develop agreed detailed guidelines and formats for data sharing and release.
- NSC under the network may select additional NSC members based on current issues when the need arises.

5.2 Membership of the network

All currently involved surveillance sites shall be members of the network. At the moment there are six university based HDSS sites (Research Centers) and one mortality surveillance program which are registered as members of the network. These sites are Addis Ababa Mortality Surveillance Program, Arba Minch Zuria Health Development Program, Butajira Rural Health Program, Dabat Research Center, Gilgel Gibe Field Research Center, Kersa Demographic Surveillance and Health Research Center, and Kilite Awlaelo Demographic and Health Surveillance Site. Except the Addis Ababa Mortality Surveillance Program, these HDSS sites

are collecting more or less similar data with similar methodology. Other upcoming HDSS sites which fulfill the membership criteria can be registered as members of the network. In fact, similar new sites are encouraged to register as member of the network since this may increase the representativeness of the data to show the country's picture. When sites are registered to become members of the network, they should agree to share their data under the premises of this policy

6.Guiding principles of sharing and releasing data

The following principles are applicable for data sharing and release:

6.1 Accountability

- EPHA shall be accountable for the network data ware house update and management
- Surveillance sites shall be accountable for assuring site level data quality, timely uploading (submitting) of data on agreed format, and data maintenance
- NSC under the network shall be accountable for proper decision based on the network data sharing and release policy

6.2 Confidentiality

EPHA, sites and NSC shall maintain confidentiality of data based on the ethical principles.

6.3 Data quality

EPHA and sites shall maintain data quality at the network and site data warehouses.

6.4 Efficiency

Releasing data to the public and sharing data with partners is an efficient way of ensuring that data are used to their full potential, that work is not duplicated. [14, 15]

7. Organizational arrangements for effective data sharing and release

The Network Scientific Committee, which is accountable for the Network of University-Based Research Centers in Ethiopia, is the main body which will decide on data sharing and release. Members of the scientific committee will be selected by the network general assembly meeting. During selection, professional diversity (e.g physician, public health professionals, EPHA legal person, demographers etc…) shall be considered. While it is important to

consider professional diversity, the network shall make sure that all HDSS sites, EPHA, CDC and the Federal Ministry of Health are represented in the committee. The Secretary of the Scientific Committee should be a permanent person hired by EPHA. The term of the selected Scientific Committee members will be two years. Meeting should be organized every quarter. As the need arises, extra ordinary meetings can be called. EPHA, through the secretary, shall take the lead to support different meetings.

Roles of the Network Scientific Committee

The Committee:

- Is accountable to the network general assembly.
- Supervises the quality of data warehouse.
- Develops its own working guideline to be endorsed by the network general meeting.
- Own data warehouse on behalf of universities.
- Accepts proposals from researchers.
- Evaluates the merit of the research by considering its relevance to the country.
- Requests clarification and, as the need arises, may request for defense of the proposal.

- Requests for ethical clearance of the research.
- Permits usage of data based on the level of access (open,restricted).
- The NSC Chair shall sign agreements with researchers.
- Cross checks the result against the proposal.
- Keeps archives of the proposals and outcomes.

The data manager at EPHA will receive data from the respective sites and store it in a database which is developed for this purpose. The respective sites shall send clean data to the EPHA data manager every year. The data manager will make sure that he/she receives a clean agreed upon data. At the end of each year, the scientific committee will analyze and produce reports. This report can be published as a scientific article or can be prepared as a policy brief for the immediate consumption by stakeholders.

8. Data sharing agreement

When an individual, group, or institution requests for data sharing, the NSC will check the request to decide whether it is possible to share the requested data. The network together with other

stakeholders must develop priority research agenda that may possibly be supported through data generated by the HDSS sites. This document will be used as a guide to make decisions when request for data sharing is submitted. If the committee accepts the request, an agreement shall be entered between the network and the prospective data user. A template/format shall be developed for Memorandum of Agreement (MOA) use. The minimum set of information in the agreement may include:

a. Name of the network

b. Address of the network

c. Name of the data user

b. Address of the data user

c. Institutional affiliation of the data user

d. Support letter from the institution where the data user is working

e. Mechanisms of maintaining confidentiality

f. Data set to be given to the user

g. Purpose for which the data will be used

h. Whether it is possible to transfer the data to third party

i. Authorship

j. Time frame the shared data will be used

9. Managing disclosure risk

Each research site and the network are responsible for minimizing or avoiding risk to individuals who provide the information. This is especially very important when sensitive data are collected from the study subjects. In such circumstances, the shared datasets must be anonymized and should contain no data elements that will allow direct identification of the subjects. The NSC will assess the disclosure risk before sharing or releasing data. Only datasets that have a low risk of identity disclosure should be released. Security systems shall be applied to access data which require restricted access.

10. Ensuring data quality

Data quality can be assured by the following strategies:

- •Standardization of the data collection tools
- •Training of research staff at all levels
- •Appropriate field supervision
- •Use of similar database for data entry

- Use of the same data entry template by the respective sites
- Regular cleaning of data at the respective sites at the end of each update round
- Sending clean data to the network based on agreed sets of variables
- Further cleaning of the data at the network level
- Sharing or releasing a clean data to prospective users

11. Network data repository and archiving

The network of network of research sites will develop a web-based catalogue which will help to list, describe and facilitate access to datasets depending on the level of access decided by the NSC. For this purpose, the iSHARE2 developed by the INDEPTH Network can be utilized. [15] The following procedures can be undertaken to share or release data.

11.1   Preparation of the web-based catalogue

- The network data manager and web-site administrator shall develop the web-page catalogue in consultation with the NSC.

• The data manager may also serve as the web-site admin istrator.

11.2 Site registration

• Each site must be registered from the network web site.

• Each site should receive a confirmation email from the net work website administrator.

11.3 Data submission by the respective sites

• Each site should regularly submit a cleaned data to the network data manager through uploading the data in the network website.

• Sites shall respect the time set for data submission to the network.

11.4 Data management at network level

11.4.1 Data downloading

• data should be downloaded by the data manager at net work level

• the data manager needs to check the quality of data im mediately and give feedback to the member sites on the completeness and accuracy of the data

11.4.2 Merging

• If possible data merging should be automatic. Otherwise the data manager at the network level should merge all the data.

• The data manager should merge data only if he/she feels he/she has complete and accurate data

11.4.3 Uploading

The network data manager (website administrator) should upload the cleaned data for access by other potential users based on the access level as determined by the NSC.

11.4.4 Data access

For all types of access, users need to be registered based on their data need. For example, if the data required is open access, they should select the domain "open access". If they request data under restricted access, they should enter agreement with the network before data is shared. If users are allowed to use the data, the granted data will be hyperlinked or ready to be downloaded for that specific user in STATA (dta) or SPSS (sav) format. Once data under restricted access is downloaded, the datasets must not be passed on to other users without the written consent of the network. All reports and publications based on the requested data

must be sent to the network Data Archive in a Portable Document Format (pdf) or a printed hard copy.

The data stored as closed access will be hyperlinked for the network data manager only. That is, when the network data manager login with his/her username and password, the data will be hyperlinked or ready to be downloaded in STATA (dta) or SPSS (sav) format.

12. Documentation of the data released or shared

There should be an appropriate documentation of the shared or released data. This documentation shall show the conditions under which the data were collected, what the data represent, the extent of data completeness and accuracy, and any potential limitations on their use. Careful documentation increases the likelihood that secondary data users will interpret data correctly. The following can be considered during documentation:

1) Name of person responsible for the data or the person to contact about using the data.

2) Overview on the data:

a) Source of the data.

b) Study design.

c) Information about data collection activities and data col
lection instruments used (e.g., a questionnaire).

d) Database (software) used for data entry

3) Period covered by the data.

4) Date and place of publication of the data (if any)

5) A data dictionary that describes the variables, data values, and coding classifications for the variables used in the original data set and for those derived from the original variables

6) A complete list of the data files that make up the data set.

7) The confidentiality procedures applied to the data in order to limit the potential for re-identification.

8) Any other information data users need, including information on limitations about the data.

9) Format in which the data are available.

10) Medium in which the data are stored (e.g., CD-ROM, Internet).

13. Data sharing and authorship

To deserve authorship, an individual should have contribution
  in one or more of the following:
 (1) Substantial contributions to conception and design, or acquisi
      tion of data, or analysis and interpretation of data
(2) Drafting the article or revising it critically for important intellec
    tual content
(3) Final approval of the version to be published.

Based on these criteria, researchers from the respective research sites and those who have scientific input at network level (E.g. researchers from EPHA & CDC) can deserve authorship in addition to those individuals who get access of the data for analysis. Generally, when agreement is entered between the network and the data users, the agreement shall include authorship. Among others, the community and funding organizations (EPHA, CDC, and possibly others who may support the network and research sites in the future) should be acknowledged in any publication.

14. Definition of terms

Data: Records of variable values collected from study units.

Data archiving: Is storing data that is no longer actively used in a separate data storage device.

Data release: Dissemination of data for public use.

Data repository: A web-based catalogue which lists, describes and facilitates access to all datasets covered by the network policy.

Data sharing: Granting access to data based on what is included in the agreement signed by the network and the individuals or organizations to whom data will be granted.

Disclosure Risk: Risk related to disclosing the data collected from the study units when different mechanisms of maintaining confidentiality are not applied.

Download: Receiving data to a local system from a remote system, or initiating such a data transfer.

Kebele: The smallest administrative unit of local government in Ethiopia

Public use data: Data available to anyone

Restricted data: Data that are shared only after agreement is signed between the network and the prospective data user.

Security: Any mechanism by which privacy and confidentiality poli-

cies are set up in computer systems.

The network: A network of Ethiopian Universities Research Centers which are registered as members.

Upload: Sending of data from a local system to a remote system such as a server or another client with the intent that the remote system should store a copy of the data being transferred, or the initiation of such a process

References

1. United Nations Statistics Division. Civil Registration Systems. United Nations. 2012

2. MEASURE DHS. Available at: www.measuredhs.com Accessed January 10, 2012

3. Federal Democratic Republic of Ethiopia, Population Census Commission. Summary and Statistical Report of the 2007 Population and Housing Census. December 2008

4. Federal Democratic Republic of Ethiopia, Ministry of Health. Health and Health Related Indicators. 2003 /2011G.C

5. Central Statistical Agency, ICF International. Ethiopia Demographic and Health Survey 2011. March 2012.

6. Berhane Y, et al. Establishing an epidemiological field labora tory in rural areas-potentials for public health research and in terventions: The Butajira Rural Health 1987-99 Program. Ethiop J Health Dev 1999. 13(Supp): p. 1-47.

7. Fantahun M, Kumbi S, Degu G, Kebede Y, Admassu M, HaileW,   Hailu S. Dabat Rural Health Project, Northwest Ethiopia, Report of the baseline survey. EJHD. 2001; 15 (special issue)

8. Tadesse T, Getachew A, Admassu M, Kebede Y, Awoke T, Me lesse T, Amsalu S, Alemu S, Andarge G, Wassie B, Tefera Y, Yifru S, Alemu K, G/Selassie S, Meseret S. Demographic and Health Survey at Dabat District in Northwest Ethiopia: Report of   the 2008 baseline survey. Ethiopian Journal of Health and Biomedical Sciences. Sept. 2011; 4(Special issue)

9. Araya T, Reniers G, Schaap A, Kebede D, Kumie A, NagelkerkeN,Coutinho R, and Sanders E. Lay diagnosis of causes of death for monitoring AIDS mortality in Addis Ababa, Ethiopia. Tropical Medicine and International Health. 2004 Jan; 9(1): 178–186.

10. Deribew A, Tessema F, Girma B. Determinants of under-five mortality in Gilgel Gibe FieldResearch Center, Southwest Ethiopia. Ethiop.J.Health Dev. 2007; 21(2):117-124

11. Gobena T, Berhane Y, Worku A. Low long-lasting insecti cide nets (LLINs) use among household members for protec tion against mosquito bite in kersa, Eastern Ethiopia. BMC Public Health. 2012; 12:914

12. Woldearegawi B, Ashebir Y, Gebeye E et al. Emerging chron ic non-communicable diseases in rural communities of Northern Ethiopia: evidence using population-based verbal au topsy method in Kilite Awlaelo surveillance site. Health Policy and Planning . 2013; 1–8

13. Arba Minch University. College of Medicine and Health Scienc   es. Available at: http://www.amu.edu.et. Accessed 10, 2013.

14. CDC.CDC/ATSDR Policy on Releasing and Sharing Data. 2005

15. INDEPTH network. INDEPTH Data Access and Sharing policy.

Editors:

Worku Sharew,EPHA

Sophia Teshome,Danya International,CDC