Peter Granda and Emily Blasczyk, 2016

Introduction

Dissemination is the process by which producers of microdata from surveys and from public and official statistics m their data available to other users. These users may include government officials, academic researchers, policymake the general public. Data may be disseminated publicly without any restrictions (as public-use files) or only to certai under specific conditions (as restricted-use files). The availability of microdata is often dependent on national laws a regulations. Data and documentation may be disseminated in various formats, but the goal should be to provide com information in a non-proprietary format that is amenable to long-term preservation.

Several aspects of making data and documentation files available to analysts require special consideration. More is i in the dissemination process than merely providing data access to interested researchers. Data producers and archivi assure analysts that the data they provide accurately reflects the efforts of the data collection process, is trustworthy fully documented, has no confidentiality concerns, and is securely preserved for future use. Disseminating data fron multinational, multicultural, or multiregional surveys, which we refer to as '3MC' surveys, can include specific proc such as standardization, harmonization, and multi-lingual documentation, which may not apply to surveys done in a country.

An additional aspect of dissemination is how to share research findings with interested parties. Determining who is data and why is important to consider as part of a comprehensive dissemination strategy. Many international organiz social science data archives, and survey research projects also embrace these objectives. Although focused on microeconomic data, the , for example, established a set of guidelines on macroeconomic data for member countries follow in order to provide the public with "comprehensive, timely, accessible, and reliable economic, financial, and demographic data" .

[↑ B](#)

Guidelines

Goal: To ensure that survey and statistical research teams in all cultures and countries involved in a 3MC survey fol accepted standards for the long-term preservation and dissemination of data to the social science research communit the wider public.

1. Make a dissemination and data preservation plan that includes archiving, publishing, and distribution, ear project lifecycle. *Rationale*

Dissemination is an integral part of the survey research process. It involves the documentation of major steps in the lifecycle, from initial planning to the production of final data files. This includes, when available and appropriate, de information about the survey process (paradata), all data editing steps, and protocols that determine which types of documentation files are made available to which users.

Procedural steps

1.1 For multi-lingual surveys, decide on the standard documentation language to be used.

1.2 Identify any documents that should be published in their original language, such as individual country questionnaires, codes, verbatim responses, and nation-specific data files.

1.3 Have a system in place to preserve all major planning and operational documents as soon as they are create

1.4 Consider including information about the survey process when disseminating data, documentation, and rep Producers may want to balance the amount of paradata they release with the need to maintain proprietary info about the data collection process.

Lessons learned

1.1 All studies must develop a system for preserving and storing materials. There are a variety of methods that utilized which rely on centralized depository. Some examples of dissemination strategies are below.

1.2 Round 4 of the strongly recommends that participating countries scan their completed paper-and-pencil questionnaires. Hard copies are acceptable where circumstances (e.g., cost) prevent scanning. National partne responsible for either the scanning or storing of their own questionnaires. Each national partner is responsible entering and cleaning their own data and delivering a clean SPSS data set.

1.3 The [Demographic and Health Surveys Program](#) (DHS) provides 'standard recode' datasets to users. The rec datasets contain the same data as the raw datasets, but in a standardized format where variable names and defi are, wherever possible, consistent across all surveys. DHS also provides a step-by-step introduction to using I data, a tabulation plan, and many other resources for analyzing DHS data.

1.4 All documents related to each round of the are uploaded to a server. This includes, but is not limited to, orig unedited (raw) data, fieldwork documents, metadata, and population statistics for coverage and response rates

1.5 Documentation of survey methods and data files is sent to a central data archive no later than nine months fieldwork is completed. Data are to be sent unweighted, but descriptions of weighting procedures should acc the datasets.

1.6 Master copies of all important [Living Standard Measurement Survey](#) (LSMS) files are kept in a separate, b archive.

1.7 Documentation for the World Mental Health (WMH) Survey is done using the Survey Metadata Document System designed by the WMH Data Collection Coordination Centre .

1.8 Countries participating in the [World Values Survey](#) are required to submit documentation of their survey m and data to a central data archive no later than three months after fieldwork has been completed. Documentati

include a completed methodology questionnaire, a report of any questions omitted or added to the original off
questionnaire, a report of additional and/or country-specific codes to any questions, official demographic stati
weights used, and a copy of the original country questionnaire.

1.9   Many institutions which provide research grants for data collection now strongly recommend that grantees
data sharing plan as part of the proposal process. The in the United States provide the following justification f
emphasis on dissemination: "data sharing promotes many goals of the NIH research endeavor. It is particularl
important for unique data that cannot be readily replicated (see [here]). Data sharing allows scientists to expedi
translation of research results into knowledge, products, and procedures to improve human health. There are r
reasons to share data from NIH-supported studies. Sharing data reinforces open scientific inquiry, encourages
diversity of analysis and opinion, promotes new research, makes possible the testing of new or alternative hyp
and methods of analysis, supports studies on data collection methods and measurement, facilitates the educati
new researchers, enables the exploration of topics not envisioned by the initial investigators, and permits the c
of new datasets when data from multiple sources are combined." This policy has resulted in more data becom
available in the public domain.

1.10   The conducted an informal Web survey of institutional data policies in the social sciences in 2013. IFDO fo
there was a growing awareness and interest in data sharing. However, the implementation varies across count
research funders. The results indicate that the social sciences have more developed policies than the medical a
health sciences.

1.11   More than ten years ago, the began to develop a set of dissemination standards "to guide countries in the pr
to the public of comprehensive, timely, accessible, and reliable economic, financial, and socio-demographic d
These standards were considered best practices, but their implementation was completely voluntary, dependin
policies and wishes of each nation. The Fund published a report about the success of this initiative over the fi
years of the initiative. It concluded that more accurate and reliable statistical information is now being produc
more nations than ever before, but also recognized that dissemination mechanisms are not fully developed in
locations. Nations also have internal challenges and constraints in addressing dissemination goals from resour
constraints, shifting priorities, and their ability to generate periodic and timely statistical data.

1.11   The has produced guidelines for access to research data from public funding. It aims to help governments, r
support and funding organizations, research institutions, and researchers themselves in dealing with challenge
improving the international access and sharing of research data.

**2. Preserve sustainable copies of all key data and documentation files produced during the data collection pr
well as those made available for secondary analyses.**

Rationale

Preservation is an important part of the survey lifecycle, a prerequisite for long-term access to valuable physical obj
digital materials. The materials that need to be preserved and kept available to members of the research community
such objects as public-use data and documentation files (including key files used in their construction), copies of the
collection instruments, user guides, information about the data collection process, and reports on field operations. Si
dissemination policies may differ among countries, it is important that data producers take the necessary steps to ma
collections as accessible as possible to members of the research community. If appropriate repositories are not avail
producers may need to organize dissemination of their materials themselves.

Procedural steps

2.1   Define the long-term preservation standards and protocols to be used. Consider digitizing physical objects,
commonly-used questionnaires, or other administrative materials documenting the whole data lifecycle, inclu

design phase of the project.

2.2   There are several digital preservation metrics that can be used to assess digital repositories. Two metrics are [Trusted Repository Audit Checklist (TRAC)](#) and the [Trusted Digital Repository Checklist (TDR)](#), or [ISO 163](#)

2.3   Protect digital materials through storage of multiple copies in multiple locations. An ideal preservation stora situation includes a minimum of several offsite copies of digital materials undergoing regularly scheduled bac it is not possible to store materials at multiple sites, preserve at least one copy in a different location.

2.4   Make certain that digital materials remain retrievable through constant refreshment of the media on which t stored. This is particularly important if removable media (i.e., tapes) are used for storage, since formats and th machines required to read these media change quickly over time.

2.5   Implement a system of version control to maintain older versions of important data and documentation files should be able to follow the changes made from one version to the next. Version control is necessary for users replicate previous analysis or to test analysis done by others.

2.6   At a minimum, store a copy of all data and metadata files in software-independent formats (e.g., ASCII or X which, with proper accompanying documentation, can be read into all major statistical packages.

2.7   Investigate the protocols and standards of digital repositories, such as availability of extracting data and in t of multi-site storage, security, and costs.

2.8   Make test runs of copied data to ensure error-free copy processes.

2.9   If possible, work with a trusted digital repository, such as a national or public social science data archive, to preserve all study materials. In doing so, data producers do their best to ensure that their data collections will available to the research community.

    2.9.1   Such repositories make an explicit commitment to preserving digital information by:

- Complying with the Open Archival Information System (OAIS) in the U.S. and similar standards in oth countries which have their own digital preservation standards and practices .
- Ensuring that digital content can be provided to users and exchanged with archives without damaging it integrity.
- Participating in the development and promotion of digital preservation community standards, practice, a research-based solutions.
- Developing a reliable, sustainable, and auditable digital preservation repository that has the flexibility t and expand.
- Managing the hardware, software, and storage media components of the digital preservation function in accordance with environmental standards, quality control specifications, and security requirements

2.10 If no national or public social science data archives exist, consider depositing data with an archive in anothe country, or investigate the possibility of doing so with a national statistical agency or certified provider. Consi archiving collections in one archive, which would keep master copies of files in several locations but minimiz possibility of conflicting versions of data and documentation files.

Lessons learned

2.1   The requires data to be archived for a minimum of 10 years as part of its anti-fraud activities.

2.2   Some earlier studies, such as older Eurobarometer surveys, did not preserve individual country data, and th about harmonization emerging some decades later could not be easily settled.

2.3   Data producers should make every effort to extract data that is on media which may no longer be easy to re
many data files have been irretrievably lost because the files were never copied to newer types of media.

3. Conduct effective disclosure analysis to protect respondent confidentiality. *Rationale*

Any plan to disseminate survey data must include very specific procedures for understanding and minimizing the ris
breaching the promise of confidentiality that is made to respondents at the time of data collection. The key goal of d
risk analysis and processing is to ensure that the data maintain the greatest potential usefulness while simultaneously
offering the strongest possible protection to the confidentiality of the individual respondents. Disclosure analysis ha
become increasingly important as more and more datasets become available online and as the possibility of linking
data to other contextual and administrative databases has grown exponentially .

Procedural steps

3.1   Be aware of, and adhere to, the different legislation for disclosure control in each country.

3.2   Disclosures can be categorized as either identity disclosure or attribute disclosure .

    3.2.1   **Identity disclosure** results from using a single identifying characteristic or a combination of characte
discover an individual respondent (e.g. name and address).

    3.2.2   **Attribute disclosure** results from using a combination of indirect characteristics associated with an i
(e.g. an individual in a specific region whose household size is an outlier and who has an unusual marita
which together reveal the individual's identity).

3.3   Implement a disclosure protocol. A proper disclosure protocol includes an analysis of the most likely outsid
sources which might allow for the identification of respondents or households.

3.4   Search systematically in the data file for sensitive information such as transcripts of open-ended answers, in
International Standard Classification of Occupations (ISCO) occupational variables, identification of primary
sampling units (PSUs), birth dates, income, or housing and dwelling information.

3.5   Search also for unusual characteristics and for cells in tables with very low frequencies.

3.6   Undertake both practical and statistical steps to identify cases and variables. This allows for the identificatio
areas or variables that need to be further masked in order to prevent identification of subjects, either through a
or by matching study data with data from other external databases. After having decided on which variables p
unacceptable risks, mask the relevant information.

3.7   Evaluate data files once those cases and variables are identified. In virtually every case, the data can be mas
various ways that make it possible for public-use data to be distributed, usually through a Web-based system.

3.8   Use appropriate masking procedures to preserve respondent confidentiality while also trying to optimize the
usefulness of the resultant data file for analysis. These procedures might include top- or bottom-coding of key
demographic variables such as income, removing data for very sensitive variables, and swapping data values
similar cases .

3.9   Document all confidentiality assurance processes and make a final assessment about the anonymity of the d

Lessons learned

3.1   With the enhanced emphasis on privacy in almost all countries, confidentiality reviews of microdata are increasingly important, if not indispensable, to assuring the future availability of public-use data.

3.2   A 2011 experiment used individual-level reoffending and sentencing data in the UK to demonstrate the poss of disclosure prior to public release. Disclosure resulted from matching data to a local news website .

3.3   The practice of reporting examples of privacy violations, particularly in the healthcare field in the United St increased awareness of this issue .

4. Consider the production of both public- and restricted-use data files. *Rationale*

In order to ensure that researchers have access to the greatest amount of data possible without compromising respon confidentiality, data producers, when appropriate, must make every effort to create both public- and restricted-use da documentation files, and make these files available to the research community through secure and predictable chann

Procedural steps

4.1   Make data files fully available to the research community as soon as possible, within the confines of how th is organized and financed. If general distribution is not feasible, establish clear rules under which researchers obtain the data.

4.2   Remain cognizant of any copyright restrictions the data may have. In some cases, even after dissemination, ownership of the data remains with the principal investigators.

4.3   Provide access directly by the data producer if resources permit, but always send copies to a trusted digital repository for permanent preservation as well, in case the data producer should cease to provide access at som in the future.

4.4   Consider the creation of less thoroughly masked versions that can be distributed under restricted-use contra made available within a research data center or 'enclave' (i.e., a secure environment in which the user has acce restricted data and analytic outputs under controlled conditions).

4.5   Establish clear policies for how researchers may access restricted data files by creating a set of application and restricted-use data agreements that specify how researchers can obtain and use such data .

4.6   Distribute restricted files through signed data use agreements. These may incorporate data protection plans, licenses, and travel to a special facility at which researchers can access the data in a very controlled environm

4.7   Create special files for researchers that cannot be matched with public-use files (for example, provide finer local information and simultaneously change respondents' IDs and other matching variables).

Lessons learned

4.1   Consider making clear agreements on data heritage (i.e., copyright transfer after the original principal inves retires). A German elite study was nearly lost to the academic public due to heritage issues.

4.2   Most data are already paid for by taxpayer money or foundations. Thus, foundations and public funders ofte for free data access (i.e., they deny the principal investigator's sole ownership on collected data).

4.3   Despite general agreement on the advantages of making data accessible to other researchers, as well as stro sharing cultures in many nations, too few social science data collections are effectively preserved. Data archiv

should do as much as possible to facilitate the deposit process by contacting principal investigators and data p
as they prepare data and documentation files.

5. Produce data files that are easy for researchers to use. *Rationale*

An effective data processing strategy focuses on the production of data files that will provide optimal utility for rese
Such files have been thoroughly checked and cleaned, possess uniform and consistent coding strategies, use commo
formats, and address the potential research needs of secondary analysts.

Procedural steps Processors should perform a series of steps to ensure the integrity and maximum utility of public-u
Such steps include:

5.1   Address the various ways data may be utilized by creating tools within a Web-based system that permits on
       analysis, subsetting, and access to documentation. Be aware that online analysis must use fully anonymized d
       users may be policymakers seeking summary information, analysts browsing for new data sources, or individu
       seeking summary analytic information or wanting to quickly download specific variables.

5.2   In order to provide optimal utility for researchers, produce a variety of products for varied constituencies.

       5.2.1   Produce setup files and ready-to-use 'portable' files in SAS, SPSS, and Stata to address the needs of tl
               seek to do intensive statistical analyses with particular software packages.

       5.2.2   Consider disseminating data on removable media (e.g., CD-ROM or DVD) if appropriate.

       5.2.3   Clearly identify the master version and provide access to any previously released versions.

5.3   Format the data files in a way that permits access through a wide variety of statistical packages, all of which
       produce the same results no matter how complicated the analysis requested, particularly with any variable wh
       decimal precision is an important consideration.

5.4   Consider creating simplified versions of datasets for use by the wider public, such as journalists and policyr
       (i.e., by creating recode variables, such as age of respondents in groups or income in groups; removing detaile
       information such as household lists; setting missing data properly; etc.). Make such datasets accessible via We
       analysis.

5.5   Make a thorough investigation of any undocumented codes or inconsistent responses, and whenever possibl
       provide labels such as 'not ascertained' if there is no alternative.

5.6   Standardize all missing data values, unless it is not possible to do so due to different cultural understanding
       such issues carefully). Users doing analyses will appreciate that all 'does not apply,' 'don't know,' 'refused,' and
       data available' responses are coded the same way in the data file.

5.7   Create complete and concise variable and value labels which will provide researchers with clear description
       their analytic results.

5.8   Provide a printable questionnaire that contains all variable names and values in an appropriate format.

5.9   Consider producing ancillary files for those data collection efforts which cover multiple waves of responder
       several geographic areas. Such files may include recoded variables to summarize information contained in ma
       questions or special constructed variables that producers feel will aid researchers in their analyses.

5.10   Create special subsets of data that take advantage of the longitudinal richness of long-term collections and p
unique opportunities to study important social, political, and economic issues from different perspectives, par
with regard to the changing characteristics of the sampled respondents. Some examples include:

   5.10.1 The project integrated a subset of data from the Demographic and Health Surveys for women of childl
age and their children from 18 countries.

   5.10.2 The created modules on specific topics that integrated data for repeating years and across countries. E
modules include 'Religion,' 'Role of Government,' and 'Leisure Time and Sports'.

5.11   Whenever possible and expedient, make individual country datasets available in 3MC surveys.

Lessons learned

5.1   Users increasingly expect data files to come in a variety of formats that will work easily with their statistica
package of choice. In some settings, this may be just an SPSS portable file, but in others, data producers and/
archives might need to create the same file in a variety of formats, particularly if a standard database conversi
package such as STAT-TRANSFER is not available.

5.2   Be very clear about coding responses that refer to 'item response refused,' 'item response does not apply due
filtering,' 'can't choose all,' or 'don't know,' and especially 'no code in data file where a code should be.' All th
different meanings and must get different values. The 'no code in data file' indicates either an interviewer erro
in data editing.

5.3   'Don't know'/'can't choose' responses may have different meanings in different countries based on different
styles. Treating all of these responses as missing data may lead to unwarranted conclusions about the attitudes
whole populations .

5.4   Established 3MC studies share their data in a variety of ways:

   5.4.1   The publicly releases all data and documentation via their website, one year after the completion of fie

   5.4.2   The releases anonymized data through their public website within one year of the onset of data collect

   5.4.3   The makes individual national and/or combined datasets available to the scientific community by the I
Archive one year after the calendar year to which it relates.

   5.4.4   Living Standard Measurement Study Survey (LSMS) data are usually available within twelve months
end of fieldwork and are published on the World Bank website for the LSMS study, as well as each coun
statistics office website.

   5.4.5   Survey of Health, Ageing, and Retirement in Europe (SHARE) data are distributed through their Rese
Data Center.

   5.4.6   The World Values Survey provides data only to participating countries for a period of two years after
fieldwork has been completed; after this period, the data are made available to the worldwide social scie
community in the form of data archives.

6. Develop finding aids to guide users in their quest to locate data collections they want to use.

Rationale

The capability to query for specific information is critical to all data dissemination systems, from individual data pro with only a few data collections to social science archives with thousands of such collections.

Procedural steps

6.1   Create a robust search engine to query the fielded metadata so that the user can find variables of interest eff

6.2   Allow the search engine to run against a study's bibliography to enable two-way linking between variables a publications based on analyses of those variables.

6.3   Display the abstracts of the publications with links to the full text whenever possible, in order to realize the potential of the online research environment.

6.4   Dedicate staff time to continuously searching journals and online databases to discover new citations where have been used. Many search engines have the ability to set up 'alerts' that notify a user when new items are f based on a query.

6.5   Encourage data archives to create metadata records for surveys they do not preserve and distribute these rec facilitate their discovery and use.

Lessons learned

6.1   Data usage increases when the data are easy to find and when users know of publications scholars have pro from the data. There are many datasets that would be of interest to secondary analysts if the analysts only kne them. For example, many surveys were conducted in Latin America and Africa in the 1960s and 1970s which offer opportunities for interesting comparative analyses with the more recent and much more popular Latino a Afrobarometer surveys. These are not always as visible to researchers, however, as they might not possess immediately obvious substantive or methodological interest.

7. Create comprehensive training, outreach, and user support programs to inform the research community al dataset. *Rationale*

Training and support of users will increase usage of the data and encourage comprehensive analyses. It is very impo that major survey research producers or archives reach out to the user community effectively, in order to explain the structure of new datasets and to encourage the greatest possible use. The most straightforward way to reach out is to an effective online presence, ensuring that the data are easily located and acquired and that metadata and bibliograpl citations are also available. Good user support will prevent obvious misuse or possible misunderstanding of the stru content of the dataset.

Procedural steps

7.1   Organize workshops at relevant professional organizations or attend conferences where 3MC research is a f soon after the data are released in order to bring early users together to discuss important preliminary results, as to ensure that the data are used effectively and that any problems with the data are recognized and correcte

7.2   Maintain a presence at professional meetings even after the data have been available for a long time. Staff f project can describe the data, distribute documentation and sample data, and encourage researchers to make u data.

7.3   Hold training workshops in different countries to ensure that novice users have a chance to learn about the experts and, if possible, from the data production team itself. Users should learn about specific issues involve

collected in their own countries, as well as how comparable the data collection experience was in other countries

    7.3.1   Without specialized instruction and training, analyses of cross-cultural longitudinal data and repeated sectional data are particularly challenging.

    7.3.2   These training courses can be brief half-day or one-day sessions at the time of professional meetings, can continue for longer periods (e.g., three- or five-day sessions with a more detailed focus).

    7.3.3   Provide the training materials online so that people who are unable to attend can still have access to th information.

7.4   Provide easy access to user support through phone, email, online chat, user forums, and tutorials.

7.5   Track all user questions in a database that creates an accumulating knowledge base and that can also serve t generate frequently asked questions (FAQs).

7.6   Create tutorials, some of which may be offered in video format, to provide help in using the data, the online system, and the major statistical software packages.

7.7   Establish moderated user forums to provide the foundation for an online community of researchers and stud who can discuss their experiences using data and learn from each other.

7.8   While all of these procedures can increase the effective use of 3MC datasets, each country must decide on v steps would be most beneficial for their own research communities.

Lessons learned

7.1   In order for participants to fully benefit from the experience, training programs must be well-planned, with level of substantive, methodological, and technical expertise. While data producers are usually the people wh understand their data, they may not have the resources or desire to provide ongoing user support for the resear community. Some may delegate this task to a data archive, but a joint approach, with data archives providing user support and data producers addressing more complicated substantive questions, often works best. In cour were national data archives do not exist, data producers may want to partner with university social science departments or research centers to increase awareness and use of important datasets.

7.2   Complex data sets often require specialized training. Data collection methods or sampling frames often cha between different waves or in different countries, and weighting variables may require extensive descriptions. context, there is no real substitute for intensive training and ongoing user support.

7.3   The Demographic and Health Surveys have an [online user forum](#) for users to post and discuss issues.

8. Produce comprehensive documentation for all public- and restricted-use data files. *Rationale*

High-quality documentation is essential for effective data use in all surveys, but particularly in 3MC datasets, becau need to provide comparable information from all countries or study populations. As resources permit, data producer strive to provide documentation, commonly referred to as metadata, on all aspects of the survey or statistical lifecyc initial planning through final data production and its release to the research community. For more information on da processing techniques used preceding dissemination, see [Data Processing and Statistical Adjustment](#).

Procedural steps

8.1   Keep detailed records from the very beginning of the project and make every attempt to record important pr
events at the time they occurred. This will assist analysts in understanding the goals and purpose of each surv

8.2   Update documentation continually during the entire lifecycle of the project, and preserve old versions of ke

8.3   For 3MC surveys, provide complete information about how the survey was conducted in each country or st
population, and describe specific procedures and practices involving data collection and data processing activ

8.4   Consider adopting the standard for producing metadata. The use of this standard, which is based on the use
Extensible Markup Language (XML), allows for specification of each metadata element (e.g., title of the surv
of the principal investigators, type of sampling) for storage and future searching.

    8.4.1   Define a database structure that will be used to store XML elements.

    8.4.2   Identify appropriate tools that will access and create XML-coded information in a natural language
        environment, such as a browser displaying a Web-based form generator.

8.5   XML metadata markup offers opportunities for data producers to create their documentation, as well as sev
advantages to users of the documentation:

    8.5.1   All information that the analyst needs is available in a core document, from which other products (suc
        files that contain the necessary information to run statistical analyses in software programs) can be produ

    8.5.2   The XML file can be viewed with Web browsers and lends itself to Web display and navigation.

    8.5.3   Because the content of each field of the documentation is tagged, the documentation can serve as the
        foundation for extraction and analysis programs, search engines, and other software agents written to ass
        research process.

    8.5.4   Preparing documentation in DDI format at the outset of a project means that the documentation will al
        suitable for archival deposit and preservation, because it will contain all of the information necessary to
        all aspects of the corresponding data files. DDI XML should ideally be generated by the CAI system use
        collect data, but can also be collected from paper-and-pencil surveys through access to the information in
        original questionnaire.

    8.5.5   There are many examples of projects that utilize DDI-compliant metadata, both at the individual study
        and in multi-study data repositories . These studies illustrate the value of using these standards, such as:

- The presentation of instrument documentation, so that users can track the logic of the questionnaire.
- The creation of question banks, comprising all items asked in multi-year studies, years items were aske
  differences in question wording, and so on. XML marked up information gets its full potential when co
  with a database management system and powerful front end tools.
- The establishment of links to the documentation of related surveys (e.g., those conducted in other count
  with variable text viewable in the native languages assists analysts who want to study relationships amo
  the survey items.

Lessons learned

8.1   Many 3MC studies provide extensive documentation online. Some examples include:

    8.1.1   The Demographic and Health Survey provides their questionnaires and manuals via their [website](#).

8.1.2   The produces an annual survey documentation report, as well as a report summarizing fieldwork and a deviations for each round.

8.2   Even though the amount of documentation that 3MC studies provide has increased in recent years, there is need to provide users with more information about the entire survey lifecycle, particularly through detailed qu profiles (see [Survey Quality](#)).

9. Consider disseminating research findings.

Rationale

Dissemination is more than storing or archiving data. Presenting research findings, in addition to making the data fil available to other users, is an important step in quality dissemination practices. This section of the chapter discusses dissemination in terms of presenting results of the study and considering who will use the information and why. Thi guideline is based on the guidelines written by the Community Advisory Board of the University of California San Francisco Center for AIDS Prevention Studies (CAB CAPS), and is adapted for the 3MC context .

Procedural steps

9.1   Create a dissemination plan.

9.1.1   Include presenting findings in the study's initial budget. This may include salary, translation, printing, and/or meeting costs (see [Tenders, Bids, and Contracts](#) and [Translation: Management and Budgeting](#)).

9.1.2   Create a team which will organize and create dissemination materials.

9.1.3   Get input from study participants, community representatives, and other potentially interested parties preferred forum for viewing findings, such as press releases, websites, newsletters, or conferences. Cons offering multiple venues, if possible.

9.1.4   Remember that there may be a need to disseminate findings several times, as new information is colle updated.

9.2   Make research results accessible to the desired audience(s). Potential audiences and effective methods inclu

9.2.1   Study participants:

- Ask participants if and how they would want to receive results. This can be incorporated as a question i survey instrument.
- Create a newsletter for participants.
- Write any information disseminated in accessible language, and keep in mind the literacy and language the study population.

9.2.2   Community members/target populations:

- Consider multiple methods, including newspaper articles, radio, and TV news, in order to reach many p
- As with study participants, consider the language needs of the community.
- Explore how research results from cross-national surveys can be disseminated to as many participating countries as possible. Different dissemination strategies may need to be employed in different countries/cultures.

9.2.3   Agencies and service providers:

- Prioritize contacting agencies that aided with participant recruitment and/or serve the target population.
- Emphasize practical use of the study results.

9.2.4   Policymakers:

- Evaluate if research results have potential to impact policy.
- Send newsletters/articles or reports to local and national government representatives.

9.3   Consider the ethical and legal policies within each country and culture. Individual countries may have diffe
dictates on sharing data within and between countries. See Ethical Considerations for further discussion.

Lessons learned

9.1   Traditionally, researchers disseminate work in peer-reviewed journals. However, practitioners, as well as th
public, rarely have the time, or even ability, to read these types of articles. The CAB CAPS guidelines were cr
a committee of activists, teachers, and other stakeholders. Committee members who had participated in resear
studies were concerned about the lack of accessible findings and developed the above points in order to addre
dissemination needs. Making the attempt to disseminate results in this way provides more benefit to those wh
the research project, and encourages discussion about the strengths and weaknesses of the original data.

9.2   The issues reports or bulletins within three months of the end of fieldwork. An advance briefing is offered t
policy makers in the executive and legislative branches of participating countries; immediately thereafter, resu
released publicly to the national and international media, civil society, and donors. Releases must be approved
core partner. Similarly, data from the World Mental Health Survey is available to policy makers in participatin
countries .

10. Make quality control an integral part of all dissemination steps. *Rationale*

The value of data depends on the quality of the data itself. Dissemination requires strict compliance to archiving, ed
publishing, and distribution protocols. Dissemination also requires the long-term availability of data and documenta
through constant updates to hardware and software and possible changes in management and staff. Clear procedures
in place to make certain all files remain readable as statistical and word processing software systems change over tir

Procedural steps

10.1   Establish a quality compliance protocol: an overall plan for regularly monitoring the integrity and validity c
and documentation files that are available for secondary use.

10.2   Consult with institutions, research associations, and analysts to develop appropriate quality standards. Stanc
should be developed with researchers to ensure they meet the needs of the relevant discipline .

10.3   Check all dissemination production steps throughout.

10.4   Test archived files periodically to verify user accessibility.

10.5   Establish procedures early in the survey lifecycle to ensure that all important files are preserved.

10.6   Create digitized versions of all project materials whenever feasible.

10.7   Develop specific procedures for assessing disclosure risk to respondents, and execute these procedures whe public-use files are produced.

10.8   Produce and implement procedures to distribute restricted-use files if applicable.

10.9   Provide data files in all the major statistical software package formats and test all content thoroughly before made available for dissemination. In addition, provide data in a non-proprietary format so that users may utili statistical package of their choice.

10.10 Designate resources to provide user support and training for secondary researchers.

10.11 Discuss with users their experiences working with the data. This may include surveying users, conference presentations, and collecting user data.

Lessons learned

10.1   The in the United States worked with other federal agencies to do a study of Web-based systems for the dissemination of health data, and produced a *Guide for Public Health Agencies Developing, Adopting, or Pur Interactive Web-based Data Dissemination Systems.* The *Guide* was developed based on the experiences of m health agencies in disseminating their data, and attempts to establish a set of general standards and practices. checklist is provided to guide agencies in developing a comprehensive Web dissemination system.

References