

# The First Nation Data Centre (FNDC) On-Site Policies and Procedures Guide



**FNIGC | CGIPN**

First Nations Information Governance Centre  
Le Centre de gouvernance de l'information des Premières Nations



# Table of Contents

- INTRODUCTION..... 4
- GETTING TO KNOW THE FIRST NATIONS DATA CENTRE..... 4
  - Orientation Session..... 5
  - The FNDC Program Manager ..... 5
- The First Nations Principles of OCAP® ..... 6
  - What is OCAP®? ..... 6
  - When was OCAP® established? ..... 6
  - Application of OCAP® to Researchers Working with Data in the FNDC..... 6
  - Are the Principles of OCAP® Especially Restrictive? ..... 10
- Confidentiality and Security Guidelines..... 11
  - Workstation security..... 11
  - File Security ..... 12
- Working with the Data..... 13
  - Tools..... 13
  - Sampling plan..... 13
  - IBM SPSS Complex Samples ..... 14
- Disclosure control guidelines ..... 14
  - Responsibilities of the Researcher ..... 16
  - Responsibilities of the FNDC Program Manager..... 18
  - Do’s and Don’ts During the Vetting Process ..... 18
- Publishing Research ..... 19
  - Providing Publication Citations to the FNDC ..... 20
  - Citing the FNDC in a Publication ..... 20
- General Administrative Information..... 20
  - Proposal ..... 20
  - Approval Process..... 21
  - Contract..... 21

Amendments..... 21

## INTRODUCTION

Welcome to the **First Nations Data Centre (FNDC)**, an initiative/project of the **First Nation Information Governance Centre (FNIGC)**. You are reading this orientation guide because your research proposal was approved by FNIGC and the Data Access Subcommittee, which means you will be granted on-site access to confidential record-level data based on the proposal attached to your contract.

The goal of the FNDC is to enable and encourage the knowledge translation of First Nations data gathered by FNIGC and its regional partners among researchers, academics, policy makers and other stakeholders. Given the sensitivity of this data, guidelines have been established to maintain the privacy and confidentiality of the respondents who have volunteered information to the data sets. To minimize risks of disclosure, all research projects will be reviewed at the proposal stage and prior to data release.

It is important to note that because of these security concerns all on-site work at the FNDC is governed by a specific set of rules and restrictions that may be atypical for many researchers. These regulations stem from the legal requirements as per the contract signed between the Principal Investigator (PI) and FNIGC. Should these rules, restrictions or terms of this contract be violated the FNIGC may:

**immediately** terminate contracts and require researchers to destroy any unpublished copies of proposed outputs -- or initiate legal actions requiring them to do so.

For these reasons, it is important that you and your research team familiarize yourself with the details of your contract and the contents of this orientation guide. Once signed, the contract is a legally binding agreement between your research team and FNIGC that allows you access to a restricted dataset based on the specifics in your proposal. If your research deviates from your proposal, requires additional data, or anything else changes within the procedures you must discuss the changes with the FNDC Program Manager and submit an amendment (details for which can be found at the end of this document).

Should the research team change at any time, you must also notify your FNDC Program Manager. All new researchers must submit signed confidentiality and acknowledgement of policies forms. If a new researcher will be physically accessing the record level data, he/she must review the on-site policies and procedures and complete the on-line OCAP® course.

## GETTING TO KNOW THE FIRST NATIONS DATA CENTRE

The FNDC is located in Ottawa, Ontario, Canada at 180 Elgin Ave. West, Suite 1200. On-site access to the FNDC is available Monday through Thursday with a lunch break between 12:30 to 1pm. The FNDC is a secure site. This means that only authorized individuals are permitted to be on site at any time, and the FNDC Program Manager (or another FNIGC employee) must be physically present at all times. You may schedule an appointment by calling 613-733-1916 (ext. 110) or e-mailing [fn dc@fnigc.ca](mailto:fn dc@fnigc.ca) . Upon arrival and leaving the office, you must sign in and out at the front desk at all times.

Any visitors will need to check-in and wait inside the reception area.

Once on site, you will work on a data access kiosk: a computer station with no internet connection, local storage capacity, or data output devices. You will also have access to a secure filing cabinet.

## **Orientation Session**

Every researcher who accesses FNIGC's record level data must first attend an orientation session that will cover on-site security protocols, the First Nations principles of OCAP®, and their contractual obligations.

The researcher will be provided with an orientation guide that details all relevant FNIGC policies related to confidentiality, privacy and security and standard operating procedures. He/she must acknowledge in writing their compliance with all of the policies and operating procedures when signing the Confidentiality Agreement.

## **The FNDC Program Manager**

The FNDC Program Manager will be your main contact regarding any questions or concerns. She/he are available at 613-733-1916 ext 110 or [fn dc@fnigc.ca](mailto:fn dc@fnigc.ca).

As the chief contact for the FNDC on-site access, the Coordinator can:

- Help potential users develop their data access proposals
- Receive, record and assess the technical merit of requests for access to the FNDC prior to approval process
- Draft data access contracts
- Schedule and coordinate usage of FNDC services among multiple individuals
- Monitor users in the FNDC
- Maintain a secure filing system for hard copies of users' analyses which users can leave with the FNDC Program Manager for safe-keeping and subsequent retrieval
- Provide methodological advice (within reason) and answer users' technical questions related to their work at the FNDC. It is, however, highly recommended that the research team include a person with experience in statistical and methodological techniques.
- Extract data sets for users based on parameters of request
- Manage users' electronic files including:
  - Uploading supplementary data;
  - Procuring, installing or arranging the installation of software;
  - Using electronic filing architecture to keep track of documentation related to users' projects, including contracts, outputs and notes
- Vet users' requested outputs by checking them for:
  - Breaches of the FNIGC's confidentiality rules;
  - Agreement with the purpose and proposed outputs identified in users' contracts
- Prepare requested outputs for release to users by:

- Suppressing information in the requested outputs as necessary;
- Producing a brief summary of the vetting process

## **The First Nations Principles of OCAP®**

### **What is OCAP®?**

The First Nations principles of OCAP® are a set of standards that establish important ground rules for how First Nations data can be collected, protected, used or shared. Standing for ownership, control, access and possession, OCAP® reflects First Nation commitments to use and share information in a way that brings benefit to the community, while minimizing any possible harm.

OCAP® means that First Nations control data collection processes in their communities, and that they own, protect and control how their information is used. Access to First Nations data is important and First Nations determine, under appropriate mandates and protocols, how access to external researchers are facilitated and respected.

It's also important to acknowledge what OCAP® isn't. OCAP® is First Nations-specific, and is not necessarily applicable to other Aboriginal groups.

OCAP® is also not a doctrine or a prescription: it respects a community to make its own decisions regarding why, how and by whom information is collected, used or shared.

### **When was OCAP® established?**

OCAP® originated in 1998 at a brainstorming session held by the National Steering Committee (NSC) of the First Nations and Inuit Regional Longitudinal Health Survey. Originally coined as "OCA," the NSC later added a "P" (in recognition of the importance of First Nations' possession of their own data) to form OCAP® as we now understand it.

In the ensuing years, the NSC became the First Nations Information Governance Committee, which in turn was incorporated into the First Nations Information Governance Centre (FNIGC) in 2010 (with a mandate from the Assembly of First Nations Chiefs in Assembly).

The principles of OCAP® are integral to the work of the FNIGC, Canada's premier source of information about First Nations people living on reserve and in northern communities. FNIGC administers two large-scale surveys, the First Nations Regional Health Survey (FNRHS) and the First Nations Early Childhood, Education and Employment Survey (FNREEES), which were designed, shaped and guided by OCAP®.

### **Application of OCAP® to Researchers Working with Data in the FNDC**

The purpose of this section is to explain how FNIGC applies the First Nations principles of OCAP® (ownership, control, access and possession) to the use of data by external researchers via the FNDC. It includes a practical set of rules that researchers should follow to ensure that their use of data adheres to the concepts and principles laid out by OCAP®.

This section is not meant to be a general treatise on the principles of OCAP® or how they apply to research. The FNIGC strongly recommends that all potential users of FNDC data familiarize themselves with the many fulsome discussions of OCAP® that exist. Some excellent papers are available on the

FNIGC website at:

[http://fnigc.ca/sites/default/files/docs/ocap\\_path\\_to\\_fn\\_information\\_governance\\_en\\_final.pdf](http://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance_en_final.pdf)

[http://fnigc.ca/sites/default/files/docs/barriers\\_and\\_levers\\_for\\_the\\_implementation\\_of\\_ocap.pdf](http://fnigc.ca/sites/default/files/docs/barriers_and_levers_for_the_implementation_of_ocap.pdf)

In addition, FNIGC has developed a video that describes the principles of OCAP®.

<http://fnigc.ca/ocap.html>

This section refers to how OCAP® applies to research proposals and data access contracts.

## Ownership

Ownership refers to the relationship of First Nations to their cultural knowledge, data, and information. This principle states that a community or group owns information collectively in the same way that an individual owns his or her personal information. To FNDC users the practical implications of this principle relate primarily to the issue of data suppression.

Most data collection agencies (including Statistics Canada) acknowledge individual-level ownership only. In theory this means that a data collection agency will not release individual-level information without the individual's permission. In practice, most data collection agencies do not release individual-level information at all.

In contrast OCAP® recognizes **four levels of ownership**: individual, First Nation (i.e. band or community), regional (i.e. treaties, nations) and national.

Like most data collection organizations, FNIGC never releases **individual-level information**.

The release of **First Nation-level data** requires the permission of each individual First Nation. Since FNIGC does not have a process in place to obtain those permissions, we do not release First Nation-level information.

The release of **regional-level information** requires the permission of FNIGC's regional partners. If you are interested in accessing this data you will have to request the regional representatives directly (FNIGC can provide you the contact information to these partners). Alternatively, you can visit <http://fnigc.ca/about-fnigc/member-organizations.html> for a complete list of regional representatives.

The release of **national-level information** requires permission from FNIGC. Permission is sought through the FNIGC's data access request forms and granted through our representative's signature on the FNIGC's data access contract.

Ownership as it relates to OCAP® should not be confused with FNIGC's legal ownership of the data we collect. Our claim to ownership of these data is typical of data collection agencies such as Statistics Canada and is described in all of our data access contracts.

Whether you are using our data centre or we send you a batch of tables, you only ever have access to the data, which should not be confused with ownership.

For researchers, the practical implications of the FNIGC's legal ownership of the data we collect are simple: use the data only in the manner described in your contract.

### *Control*

The principle of control affirms that First Nations, their communities and representative bodies are within their rights in seeking to control over all aspects of research and information management processes that impact them. First Nations' control of research can include all stages of a particular research project-from start to finish. The principle extends to the control of resources and review processes, the planning process, management of the information and so on.

This element of OCAP® seems to have caused confusion and concern among some in the research community. Some misinterpret "control" to mean that FNIGC will bury findings it does not like or will prevent researchers from expressing their true opinions of the causes of and potential solutions to challenges that exist on reserve communities. This is not the manner in which FNIGC exercises the principle of control. With respect to the use of FNDC data by external researchers, we exercise this principle through our research proposal/contract system.

The review process allows us to understand the purpose, rationale and desired outcomes of your study. These processes allow us to weigh the risks and benefits of potential findings (including findings that may mislead or be misused) and to decide in advance whether your study is something the FNIGC should support by providing you with access to the data you request.

Your signature on your contract ensures that you will use the FNDC data to which you are granted access only for the purposes approved by the FNIGC.

In practical terms, this means that you may address only the research questions described in your contract using only the methods described in your contract. If you propose to look for a link between self-reported health and smoking, do not investigate the link between gender and diabetes, even if the data to which you have access permit you to do so. If you propose to test for a specific interaction effect in your regression model, do that. (Amendments to methodology are of course possible with written permission from the FNIGC).

Once your analyses are complete, you are of course free to interpret your results as you see fit and to distribute your research in the manner described in your contract.

It is important to mention that, while we will never find a researcher to be in breach of contract simply for interpreting data in ways "we do not like," we are by no means indifferent to those interpretations. Should a researcher's analyses of FNDC data prove unreasonably biased or contrary to the interests of First Nations, we may invoke our right of dissent and criticize the work publically through editorial comments or academic peer review. We may also decline to enter into future contracts with the researcher. To avoid this situation we advise researchers to keep in mind a foundational principle of OCAP® and FNIGC: that research about First Nations should never harm First Nations or First Nation communities. If you are concerned that your research may do harm, we encourage you to talk to the

FNIGC representative who is handling your contract. We will offer advice on how to best interpret patterns and on the most careful and respectful way of reporting contentious findings.

Under exceptional circumstances, the FNIGC may decide that access to the FNDC data must be conditional not only on their approval of the proposal, but on their review of the materials the project produces. Exactly what this added dimension of control will entail will vary from project to project. If the FNIGC does require this additional dimension of control, however, it is important that the researchers affected keep two things in mind. First, this approach represents a compromise on the part of the FNIGC, which is sufficiently wary of a proposal to avoid approving it outright, but nonetheless wants to facilitate useful research whenever it can. Second, all conditions of data access for a specific project will be detailed in the data access contract associated with that project. If those conditions seem too cumbersome or risky to researchers, they do not have to enter into the contract.

In a nutshell, the proposal/contract approach with which the FNIGC “controls” the use of FNDC data ensures complete transparency (i.e. researchers know what they must and must not do) and predictability (i.e. researchers know exactly what the outputs of their projects will be and exactly what risks of censorship, if any, they may face). All researchers must do is adhere to the terms of their contracts.

### *Access*

First Nations must have access to information and data about themselves and their communities, regardless of where it is currently held. The principle also refers to the right of First Nations communities and organizations to manage and make decisions regarding access to their collective information. This may be achieved, in practice, through standardized, formal protocols.

Accordingly FNIGC has established a system to facilitate access to all information derived from the FNDC. Contracts require researchers to share a copy of their research products with the FNIGC, which we keep in a library that First Nations survey participants can access. 

The library also includes all outputs (e.g. tables) that have been released to clients.

### *Possession*

While ownership identifies the relationship between a people and their information in principle, possession or stewardship is more concrete. It refers to the physical control of data. Possession is a mechanism by which ownership can be asserted and protected. All raw data gathered by FNIGC are located on servers located within the FNIGC offices. That it must not be removed is asserted in contracts governing access to the First Nations Data Centre.

The FNIGC exercises the principle of possession in another way, however. The data access contracts assert that each table or other output can be used by only specific people for specific purposes. To ensure compliance with this provision, **each output released to a specific client for a specific purpose has its own unique name**. Therefore a person who has not been authorized to use this output cannot cite it. Since using an unreferenced output will lack legitimacy, the prospective user is therefore compelled to access the table through the FNIGC.

To FNDC users, therefore, the principle of Possession means that there are no shortcuts to accessing FNDC data. Even when those data appear to be in the public domain, mechanisms are in place to ensure the FNIGC continues, for all practical purposes, to possess them. The intent is to ensure that those wishing to access FNDC data must always do so through the FNIGC, respecting the processes we have put in place to safeguard them.

Exceptions are published data. Data access contracts permit researchers to include limited amounts of FNDC data in their publications. Anyone wishing to reproduce those data, of course, has only to cite the publications in which they appear.

### **Are the Principles of OCAP® Especially Restrictive?**

Some may be wary of seeking access to FNDC data owing to the perception that the rules governing their use are overly restrictive. In reality, however, most of those rules are standard in the field of social science research and those that are not remain in keeping with the field's accepted ethical standards.

Most of the differences between the FNIGC's FNDC data access process and those employed by other data collection agencies (e.g. Statistics Canada) hinge on the Ownership aspect of OCAP®. That is, they hinge on the fact that the FNIGC recognizes collective as well as individual ownership and its corollary, the right to privacy.

While the importance of collective rights has strong roots in First Nations cultures, the need to protect the privacy of First Nations collectives also reflects the fundamental ethical principles of mainstream social scientific research.

In Canada, most social research is governed by the *Tri-Council Statement on Research Involving Human Subjects*. It notes that individuals' right to privacy is rooted not in some abstract notion of the supremacy of the individual (as opposed to the collectives to which they belong), but in the belief that individual-level information is uniquely capable of harming the informer. It follows that, where codes of research ethics ignore the importance of collective privacy, they simply do not perceive aggregated data as capable of causing harm to the aggregates or their members.

The capacity for community- and regional-level data to harm First Nations, however, is significant. Those data can cultivate negative stereotypes that can result in discrimination against First Nations peoples. Perhaps more importantly, however, those data can be used to inform the decisions of the non-First Nations governments who wield an extraordinary amount of power over First Nations' resources.

Essentially, when First Nations are the subject of research, the extension of information ownership and privacy rights to First Nation collectives is necessary to actualize the spirit and intent of the ethics code that governs most social research in this country.

Once this need is acknowledged, it becomes obvious that the other provisions of OCAP® are no more arduous than those described in the *Tri-Council Statement on Research Involving Human Subjects*. Most significantly, researchers using human subjects are generally expected to submit a project proposal to a

review board, and to conduct their research in the manner approved by that board. This approvals process, like the FNIGC's, focuses on ensuring research rigor and subjects' protection from harm.

It is true that the *Tri-Council Statement* exempts secondary data from ethical review if it is not "identifiable" - i.e. it cannot identify individuals and thus cannot render them vulnerable to harm. However – as was explained above – community-, regional- and even national-level First Nations data can reasonably be regarded as "identifiable."

The provisions the FNIGC put in place ensures that access to FNDC research are likewise typical of data collection agencies. Tables produced for clients are maintained as a matter of course, and institutions such as Statistics Canada require researchers to report on research conducted in their Research Data Centres. Indeed, the FNIGC's requirements are arguably less cumbersome than Statistics Canada's, since the latter requires a report even if the research yields no interesting results. The FNIGC only seeks copies of reports intended for distribution. On the other hand, the FNIGC requires copies of research based on tabular data while Statistics Canada does not. The ultimate point is that most not-for-profit data collection agencies demand various sorts of "returns" from researchers in exchange for data access.

Ultimately, it may be argued the principles of OCAP® apply to most data. Every data collection agency, every organization that collects any kind of information is allowed, within limits, to decide how and by whom that information can be used. In the case of FNDC data, OCAP® lies with the FNIGC, and through their collaborative processes, with their First Nations partners.

## **Confidentiality and Security Guidelines**

### **Workstation security**

The FNDC is a secure environment with its own, isolated computer network that has no internet access, local storage capacity or enabled output devices (i.e. the ability to download information). All research files and data are stored on the FNDC network.

In addition no printers are directly attached to the workstations. Any printing will be done on a network printer under the supervision of the FNDC Program Manager using coloured paper. All printed or written information that is produced on coloured paper at the FNDC must remain on-site.

You will have a file that will be stored in a locked file cabinet at the FNDC. **A researcher working in the FNDC must not operate any personal computing device (e.g. laptop computers, PDAs, etc.) on-site.** A cell phone (or similar communication device) should not be operated in the FNDC and must never be used to communicate/discuss confidential data or outputs with team members.

Researchers can leave their cell phones on vibrate mode, but any cell phone calls must be conducted outside the secure workstation area. A portable phone in the office will be made accessible should you need to seek further assistance from your research team.

The computer equipment provided for use must never be removed from the premises of the FNDC.

At the end of the work session you must log off. If away from the workstation for a short break, you should lock the computer to prevent others from accessing it.

All FNDC computers are shut down at the end of each work day. The FNDC Program Manager may force a shut down even if you are still logged on to the network, which means all unsaved work could potentially be lost. Permission is needed to run programs overnight.

In an emergency, such as a fire, you are still expected to lock, log off or shut down the computer terminal before leaving the FNDC.

### **File Security**

You must keep their project username and password confidential at all times. The username and password is specific for the project and not for the user, consequently only one researcher can work in the project folder at any time.

When a researcher is working on multiple projects separate project folders will be created with separate usernames and passwords. A researcher cannot be logged into more than one project at a time and is not permitted to transfer files between project folders. **At no time should you change the name of the project folder.**

Projects will have the following naming convention: YYYY (year)-LASF (first three letters of PI's last name and first letter of PI's first name)-000(project number). For example, 2013-SANM-001.

Within the project folder, there will be three other folders: **Working, To Be Vetted, and Final**. The requested dataset and any previously prepared syntax will be saved in the **Working** folder by the FNDC Program Manager. The researcher's access privileges include creating, changing and deleting files/folders within this Working folder. The **To be Vetted** folder will include any output and syntax that the user would like the FNDC Program Manager to screen for release. The researcher has no access privileges in this folder, except to insert and read files within this folder. The **Final** folder will include the signed contract and confidentiality agreements, any released syntax and output, a copy of the requested and final dataset. The researcher will only have viewing access to the Final folder.

Only the FNDC Program Manager can insert external files into the project folder. It is advisable that you send via e-mail any material (including codebooks, statistical reference materials, programs/documentation etc) that you wish in your folder in advance so that it can be reviewed and inserted into your folder by the FNDC Program Manager. A researcher cannot install software on the FNDC computer.

Linking the record level data file to any other files in order to relate the particulars to any identifiable individual person or other entity is not permitted. You **cannot** remove any of the original data set or

copies of subsets of the record level data file or any confidential sensitive statistical information provided.

The FNDC Program Manager reviews and loads to, or extracts from, the server all files entering or exiting the FNDC. Therefore, you should make arrangements with the FNDC Program Manager, in advance, to reduce wait time.

The researcher **cannot** remove any data or analyses (in electronic or paper format) from the FNDC, including outputs that have been released by the FNIGC. The FNIGC will transfer encrypted electronic copies of released outputs to the Investigator(s) by email or other electronic file transfer medium, or by courier. The FNDC reserves the right to search all materials leaving the facility.

FNIGC will ensure up-to-date virus protection of all files stored in the FNDC. FNIGC backs up all files on the FNDC network on a daily basis.

## **Working with the Data**

### **Tools**

To assist you while working with data, the FNDC has created some tools designed for use by the researcher. These include:

- a data dictionary,
- the questionnaire,
- information on data collection, sampling and survey design, and
- the *Complex Samples User Manual*, which includes instructions on how to use a plan file.

These will be saved in your project's Working folder.

In addition, the FNDC computer environment includes the following applications: SPSS Base and Complex Samples, MS Word and Excel.

The use of syntax files to construct data sets and generate output is strongly encouraged. Maintaining syntax files is important to save time when similar analyses are conducted with a different variable or sub-population, or if there is a re-release of data and analyses need to be re-run. In many cases, if the same data set is to be used for a new FNDC contract, the syntax can be transferred to the new project.

Furthermore, syntax files are requested by the FNDC Program Manager as supporting documentation when vetting outputs for release.

### **Sampling plan**

The First Nations' RHS sample design incorporated a two-stage sampling strategy; the first stage involved the selection of communities to participate in the survey. First Nations communities were

stratified by region, sub-region, and community size: large (population 1,500 + ), medium (population 300 - 1499 people), and small (population < 300). Communities with a population of less than 75 were excluded from the sampling frame. Large communities were automatically included, while medium and small communities were randomly selected with equal probability within their respective strata.

The second stage of sampling pertained to the selection of individuals within each community sampled. Community members were identified using band membership lists. Data were gathered to represent eight categories of the community population (gender by four age-groups). The sampling rate within each community was determined as a function of the overall sub-region probability (within regions) and the probability of selection of the community (within sub-region).

### **IBM SPSS Complex Samples**

As a result of the above sampling approach, estimates derived based on the assumption that the data arose from a simple random sample would be technically biased. For instance, estimated standard errors of statistics would often be too small, giving the user a false sense of precision. IBM SPSS Complex Samples incorporates the sample design into the analysis, subsequently enabling the user to achieve more statistically valid estimates for the population of study.

IBM SPSS Complex Samples uses something called a **plan file** which contains the information needed by analysis procedures to properly compute variance estimates for complex samples. The plan includes the sample structure, estimation methods for each stage, and references to required variables, such as sample weights.

The plan file must be engaged prior to performing any type of analysis in the complex samples module.

### **Disclosure control guidelines**

The **disclosure control guidelines** are governed by the principles whereby the confidentiality of respondents are assured and remain anonymous when the results are reported. Disclosure control refers to the measures taken to protect data so as not to violate confidentiality requirements. The goal is to maintain confidentiality while upholding the usefulness of the data outputs.

All data and results to be released from the FNDC are carefully screened to ensure that there is no risk of disclosure of confidential information or information that may lead to the identification of an individual person, regions, community or First Nation band. You are responsible for taking all precautions to avoid disclosure of confidential information. If you discover or can inadvertently deduce small cells, regional information or any individual-level information, do not disclose that information with anyone or in any publication. In addition, we ask that you immediately bring it to the attention of the FNDC Program Manager.

Record-level data in any form **can never leave the FNDC**. This also applies to any approved non-FNDC record-level data or aggregate data brought in by researchers to be used in conjunction with the confidential data housed in the FNDC.

Output produced, used and stored **within** the FNDC will not be subject to the vetting process. When a researcher requests to remove output or syntax from the FNDC, the vetting rules are applied. All outputs and syntax leaving the FNDC must go through the vetting process, first by the researcher and then by the FNDC Program Manager.

Files in the FNDC will either be electronic and stored on the FNDC server, printed on **coloured** paper and stored in a locked filing cabinet at the end of each day, or shredded (when no longer needed).

Any hand-written notes produced in the FNDC will be written on **coloured** paper. At the end of a work session, these notes are to be given to the FNDC Program Manager for either shredding, storage, or vetting (if the notes contain no confidential information and the researcher wishes to remove them from the FNDC).

Unvetted material in electronic, printed or handwritten form is presumed confidential and **cannot** be removed from the FNDC.

Researchers are permitted to bring their own supporting documentation inside the FNDC and researchers can leave it on-site. However, this outside material should not be printed on coloured paper to avoid confusion with the confidential material.

Although direct identifiers (name, address, telephone numbers) which provide explicit information about or link to a respondent's identity are removed from FNIGC's record level data, most indirect identifiers (age, sex, occupation) are not. When indirect identifiers are used together, they can potentially lead to disclosure. This can occur when confidential information which can be attributed to an individual is revealed. It is not necessary for a specific individual to be identified or for a specific value to be given for attribute disclosure to occur.

For example, publishing a narrow salary range for persons in a specific profession living in a region could potentially constitute a disclosure. This is an issue when researchers are using narrowly defined geography or easily distinguishable population groups for their analyses.

In the FNDC, output generally refers to any statistical result produced by analyzing the confidential record level data. These can include but not limited to frequency tables, univariate statistics, or estimates of regression model parameters.

Outputs, including aggregated frequencies, means, modes, maximum or minimum values, proportions, percentiles, prevalence and ratios used to summarize or describe a sample, are generally referred to as **descriptive statistics**. These statistics pose a greater risk of disclosing the identity or an attribute of a respondent. Variance estimation and inferential statistics of model outputs such as coefficients of variation, F-test statistics, estimated coefficient values, maximum likelihood estimation, etc. pose less risk of disclosure.

Information based on an individual survey record is data and is not considered to be statistical output. This can also be viewed as anecdotal information and **cannot** be removed from the FNDC.

Using sampling weights in the data analysis helps protect respondent confidentiality. This is because when the weights are applied each respondent in the sample represents a certain number of persons in the target population. When descriptive statistics and tabulations have been weighted, there is far less chance that information specific to one respondent may be isolated and attributed to an individual. This is because the weighted frequency is no longer a raw count of respondents in the sample.

In general, output released from the FNDC must be produced using weights (generally referred to as weighted output). Results using raw data (generally referred to as un-weighted counts) can be released, but are not recommended for publication or use in presentations. Since these un-weighted estimates are influenced by the survey design, the population inferences may not be valid.

## **Responsibilities of the Researcher**

**Intermediate output** may be produced during the exploratory and model development stages of the analysis. These outputs often consist of detailed tables to check for sample size and data quality, and different ways of recoding a variable. Such outputs often present more detailed information than what is normally found in publications. The release of such detailed, overlapping results can greatly increase the risk of disclosure. Instead, these discussions should remain within the FNDC. Team members are encouraged to meet in the FNDC to discuss these results or have a teleconference on a secure, land line.

The vetting process requires time and is conditional on the number of active contracts in the RDC and the number of vetting requests. The RDC Coordinator will attempt to process a vetting request in a timely manner within a maximum of 2 weeks. However, researchers are to account for the processing time when planning research activities and submitting output for release.

If a researcher suspects a potential disclosure risk in their outputs they should discuss their results with the FNDC Program Manager as soon as possible.

When requesting for the release of output, researchers must insert the files in the To Be Vetted folder and provide the FNDC Program Manager with the extent of possible residual disclosure. Residual disclosure occurs when released information can be combined with other available information to obtain confidential information. For example: while a table alone may not qualify as a risk for disclosing confidential information, it could occur by combining information from several sources (e.g., suppressed data in one table can be derived from other tables).

To help the FNDC Program Manager understand the output, researchers should provide a description of the derived variables if the syntax does not provide sufficient information for vetting.

**Note:** The FNDC Program Manager may review data sets and syntax files used to produce the final releasable output.

For descriptive results, researchers must provide weighted (by applying the survey weights) and un-weighted (outputs from the raw data) data along with the coefficients of variation (when applicable). FNDC Program Managers rely on the un-weighted outputs to confirm that the results meet the

thresholds of the minimum number of respondents necessary to produce releasable estimates. After vetting is completed, the outputs are released from the FNDC to the researcher.

Before releasing or publishing any estimates from the data files, the researcher must first determine the number of sampled respondents having the characteristic of interest to ensure that enough observations are available to calculate a quality estimate. For example, this would be the number of respondents who smoke when interested in the proportion of smokers for a given population. For users of the FNDC, it is recommended to have at least five observations in the un-weighted numerator. Otherwise, neither the un-weighted nor weighted estimate will be released regardless of the value of the coefficient of variation for this estimate. For weighted estimates, based on sample sizes of five or more, researchers should determine the coefficient of variation of the estimate and follow the guidelines below.

### Sampling variability guidelines<sup>1</sup>

Type of Estimate	CV (%)	Guidelines
Acceptable	$0.0 \leq CV \leq 16.5$	Estimates can be considered for general unrestricted release. Requires no special notation.
Marginal	$16.6 \leq CV \leq 33.3$	Estimates can be considered for general unrestricted release but should be accompanied by a warning cautioning subsequent users of the high sampling variability associated with the estimates. Such estimates should be identified by the letter E (or in some other similar fashion).
Unacceptable	$CV > 33.3$	FNIGC does not release estimates of unacceptable quality. Suppressed values should be flagged with the letter F and a footnote stating that the F statistic represents a suppressed value due to low cell count ( $n < 5$ ) or very high sampling variability ( $CV > 33.3$ ).

<sup>1</sup> Derived from Statistics Canada, Canadian Community Health Survey (Annual component)- User guide 2010 and 2009-2010 Micro-data files, [http://www23.statcan.gc.ca/imdb-bmdi/pub/document/3226\\_D7\\_T9\\_V8-eng.htm#a11](http://www23.statcan.gc.ca/imdb-bmdi/pub/document/3226_D7_T9_V8-eng.htm#a11) (Accessed August 8, 2013)

## **Responsibilities of the FNDC Program Manager**

The FNDC Program Manager's primary responsibility in the vetting process is to ensure confidentiality is not breached when allowing research outputs to leave the FNDC. The FNDC Program Manager reviews all the materials that the researcher would like to remove from the FNDC. The final responsibility and decision to release the output rests with the FNDC Program Manager and FNIGC's Senior Manager for Research and Information Management.

In the event a disclosure risk is identified, all attempts are made by the FNDC Program Manager, in collaboration with the researchers, to find a solution. However, the FNDC Program Manager must maintain confidentiality of survey respondents and when the risk of disclosure is too high, the output will not be released.

The vetting process usually takes two weeks to complete, but is contingent on the number of active contracts ongoing at the FNDC, the number of vetting requests and the size of the output.

## **Do's and Don'ts During the Vetting Process**

Here is a simple list of suggestions for what a researcher can do (and shouldn't do) when it comes to the vetting process.

### **Do**

- Schedule time with the FNDC Program Manager advance of an initial vetting request; and well before any anticipated deadlines.
- Remove any output that you feel could lead to the identification of any First Nation, community, region or individual.
- Remove any extreme values or values representing an individual from your outputs. (Examples include minima, maxima, medians, and modes. If a procedure creates extreme observations, 0, 1, 99 and 100 percentiles, those extreme values must also be removed.)
- Mark any cell with a frequency less than five via highlight, asterisks or circled, before they are submitted to the FNDC Program Manager. If at all possible we encourage you to consider re-categorizing the variables instead of requesting suppression.
- Make a special request if you wish to have intermediate output released, though we discourage its release it may be used on-site. (Examples of intermediate output include: tables of preliminary descriptive statistics and large numbers of preliminary regression models.) The reason intermediate output increases disclosure risks is that similar tables based on different sub-samples may cause complementary disclosure problems. Comparison of tables could reveal information about the sample and individual characteristics.
- Provide weighted and corresponding un-weighted output as well as coefficients of variation for all descriptive statistic outputs.
- Restrict cross-tabular analysis to two or three dimensions.
- Limit the release of tabular output to just before the end of the project.

- Double-check for any output errors and remove needless results before submitting the files for vetting.
- Provide all necessary details to the FNDC Program Manager for vetting. If applicable, inform the FNDC Program Manager if similar output has been previously released (include the details of the differences between similar requests).

**Don't**

- Request tables for which the corresponding un-weighted cell counts are below five.
- Request anecdotal information (such as minimum/maximum statistics) or record-level data. You can look for outliers but do not include this information in your final output.
- Request graphs with individual points plotted or that show outliers (e.g. scatter plot of the raw data or box-plots).
- Printing individual cases is not permitted. Remove any individual level data from your output.

**Publishing Research**

It's important before you **publish** your research that you:

- not reveal any information that could identify an individual or First Nation band or a region;
- not reveal any information that could be used to identify geographic areas where respondents live or were sampled unless you specifically have permission to make estimates for those areas ;
- not reveal any information about specific dates
- ask the FNDC staff for help any time you have a disclosure concern

Before you submit your research for publication it's important that you ask yourself these questions.<sup>2</sup>

Question	Action
<i>Did I mention anything in the publication that I learned while looking at restricted data?</i>	If you are discussing something that you observed while in the FNDC, but did not specifically include in output for review, share this information with your FNDC Program Manager to assess if it poses additional risk.
<i>Did I use restricted geographic variables in my analysis?</i>	Remove any specific geographic information (e.g. province/territory, region). If you are discussing the number of geographic units represented and didn't specifically state that you would in the

<sup>2</sup> Derived from the National Center for Health Statistics Research Data Center, Disclosure Manual- Preventing Disclosure: Rules for Researchers, <http://www.cdc.gov/rdc/Data/B4/DisclosureManual.pdf> (Accessed July 22, 2013)

	approved proposal, ask your FNDC Program Manager to review the statement.
<i>Does my data set include any restricted temporal variables?</i>	If you want to mention aggregated dates (e.g. season, year), this must have been explained in your proposal.
<i>Could an individual or First Nation band/community/region possibly be identified by any information in my publication?</i>	Be mindful that the way you discuss inclusion and exclusion criteria could inadvertently identify a small cell or extreme case.

***Providing Publication Citations to the FNDC***

Copies of materials for all publications, working papers, presentations, and reports that refer to research conducted in the FNDC must be emailed to [FNDC@fnigc.ca](mailto:FNDC@fnigc.ca) 30 days before the day they are made available to the general public. Copies of materials that will not be made available to the general public (e.g. a presentation) must be provided to the FNIGC in electronic form a week prior to distribution to their intended audience or segment thereof.

***Citing the FNDC in a Publication***

The researcher(s) must cite the source of the released output in all documents containing the output or extracts or analyses. Citations should take the following form: Source (or 'Adapted from,' if appropriate): The First Nations Information Governance Centre, (insert specific name of record level data extract), (insert reference date of record level data extract). This citation should appear below tables or figures containing outputs or extracts or analyses thereof. If outputs or extracts or analyses appear within the textual body of a document, the citation may appear alongside them in parentheses, or may be attached to them by way of a reference to a footnote or endnote.

The researcher(s) must display prominently the following notices in all documents that contain released outputs or extracts or analyses: 1) “Analyses contained herein of data from the First Nations Regional Health Survey do not necessarily reflect the views of the First Nations Information Governance Centre (FNIGC)”; 2) “Statistics reproduced from this document must be accompanied by a citation of this document, including a reference to the page on which the statistic in question appears.”

Please avoid statements such as:

- Analyses of these data were approved by the First Nation Information Governance Centre.
- The First Nation Information Governance Centre reviewed and approved our results.

**General Administrative Information**

***Proposal***

When reviewing a data request, the criteria examined include but are not limited to:

- Demonstrated reasonable need for the data to answer the specified research question. This includes ensuring data is not accessible via the public use tables on the FNIGC website.
- Research is not harmful to the respondents or community.
- Benefits are clearly in the First Nations interest.
- Methodological and technical review
- Ethical considerations
- Not proprietary research such as research done for commercial marketing purposes.

### *Approval Process*

In the first phase, the FNDC Program Manager will review the technical aspects of the data access proposal. If modifications are required at this stage, the FNDC Program Manager will advise the researcher to make the changes and resubmit. If approved, the proposal will continue onto the second phase.

In the second phase, the FNDC Program Manager and Senior Manager for Research and Information Management (or Senior Manager for Methodology and Statistics) conducts a methodological and an ethical review of the proposal, suggesting modifications if necessary. If major modifications are required at this stage, the FNDC Program Manager will advise the researcher to make any changes and resubmit. The proposal will then continue onto the third phase. It is important to note that continuing onto the third phase does not guarantee final approval. The intent of the initial phases is to prepare the proposal for final approval by the FNIGC Data Access Subcommittee.

In the third phase, the FNIGC's Ottawa office will provide a summary of the phase one and two reviews to the FNIGC Data Access Subcommittee, along with a copy of the proposal, and any other supporting documentation.

### *Contract*

Please review your signed contract prior to working with the data at the FNDC.

Only those members of the research team who will either need access to the data or need to assist another team member accessing the data should be included on the contract. Any discussions regarding preliminary research results can only be held among those on the contract who have signed the confidentiality agreement. Help on software or statistical methods should come from members of the research team listed on the contract.

### *Amendments*

Changes to the research project may require an amendment application that must be reviewed and approved by FNIGC and confirmed with a new signature to the Research Agreement. Such changes include, but are not limited to:

- Changes to the research question(s)
- Changes to Data fields requested

- Changes to years of Data requested
- Changes to cohort definition
- Extension of the contract
- Addition or removal of members in the research team