

# Providing Context for Understanding: Insight from Research on Two Canadian Health Surveys

## Introduction

A significant question for data producers, as well as for the IASSIST community, is whether today's data documentation preserved in the form of user guides and codebooks has all of the information necessary for the analysis of a survey. This is particularly important as the community wrestles with the adoption and implementation of Data Documentation Initiative (DDI) projects to describe data from many of our national surveys.

To examine this question, I used a case study based on major population health data in Canada and employed a life-cycle perspective. I found that, with respect to the National Population Health Survey (NPHS) and, more recently, the Canadian Community Health Survey (CCHS), much of the important information and research related to placing the survey data in context is derived from activities that precede extensive data analysis.<sup>2</sup> It is the argument of this paper that the work involved in the creation of a data collection, which occurs early in the data life cycle, is integral contextual information and as such should be identified, described and preserved, in addition to the formal data collection itself.

## Background

Launched in 1994-1995, the National Population Health Survey is significant in that it was Statistics Canada's first national longitudinal health survey designed to fill a specific and critical data gap in health information: the determinants of the health of Canadians over time.<sup>3</sup> The problems related to health data were extreme from the perspective of researchers and policymakers alike. In a summit on health information sources, participants enumerated the problems with existing data sources.<sup>4</sup> While numerous statements revealed frustrations with the inadequacy of the state of health information, some critics went further and described the state of health information of the era as being in a deplorable state, suffering from a lack of comparability and with serious gaps in coverage.

The NPHS was a key component in the new health information infrastructure for Canada, and it was the hope of Statistics Canada that flowing from this survey would be the development of health indicators akin to economic

by *Elizabeth Hamilton*<sup>1</sup>

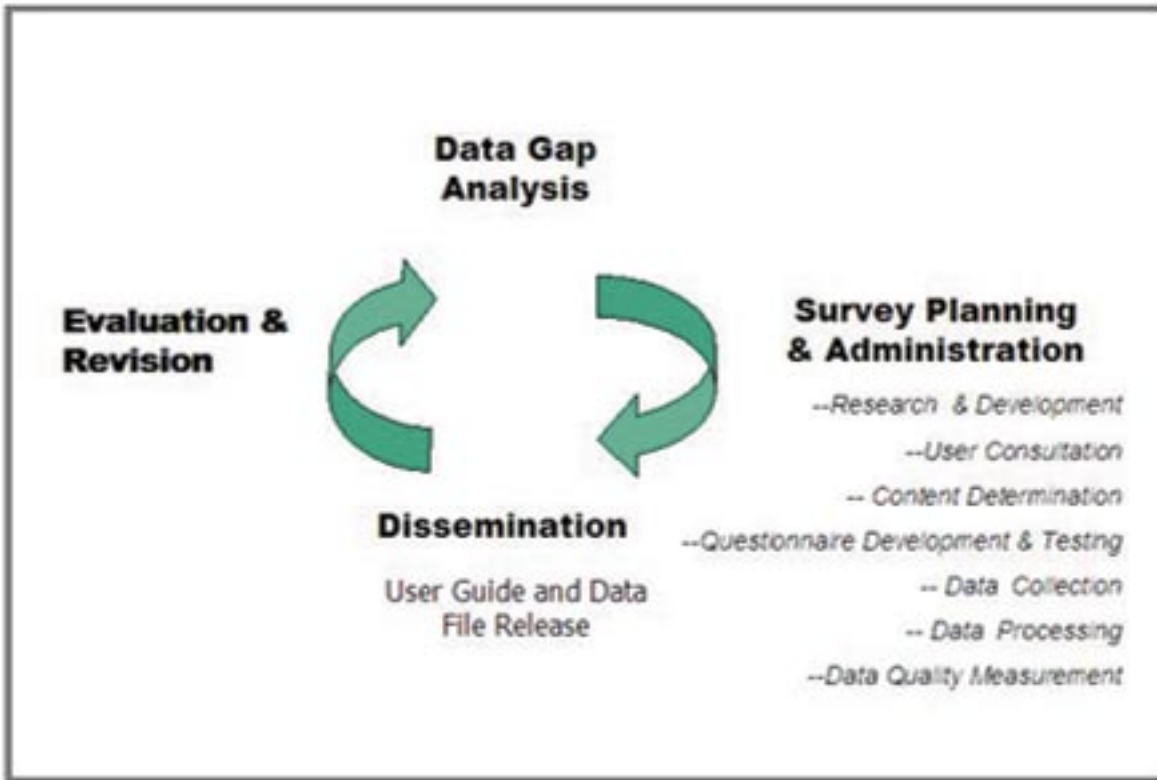
indicators for the country, a theme that has been echoed by many in the health field since the inception of the survey. Not surprisingly, the expectations for this survey were enormous. By the Spring of 1992, Statistics Canada had Treasury Board's support of the project, and research and consultations on the design and methodology of this new longitudinal

health survey were well under way. The primary survey instrument was finished by Fall 1992, a time frame that, in the best of circumstances, did not allow for the luxury of repeated revisions.

The pressure was considerable. The project managers had to devise an instrument that would fill the data gaps over more than just one survey cycle; they had to create the appropriate content, questions, and scales to collect person-oriented health information and meet the needs of researchers and policymakers over a 20-year period. And, as with many national surveys in Canada, the sampling was complex to accommodate political and social realities in the country. Additionally, companion surveys to the primary household survey were developed for institutions and for the traditionally under-surveyed northern areas of Canada.

The process of launching the survey came with guidelines that emanated from the Task Force on Health Information. The survey was to be flexible and statistically reliable, with timely release of data. Further, the process was to be consultative, allow for supplementary content or sample size, and permit linkage with administrative data. The guiding force came from the project team within Statistics Canada with oversight by a Federal-Provincial-Territorial Committee. The project manager brought together expert groups of six to ten people as needed to decide upon about ten minutes of questions to capture data necessary on mental health, health measurements, and other key content areas.<sup>5</sup> The questions were brought to focus groups and the questionnaire was modified for field-testing. The sampling design was drawn up, modified, and re-modified.

Linkage with administrative data was a key component of the NPHS, and in the survey respondents were asked if they would permit information to be shared with provincial health departments, Health Canada, and Human Resources Development Canada (HRDC) for statistical purposes. To



**Diagram 1**

solve the dual problems of confidentiality and researcher access, the NPHS team produced three cross-sectional files, plus a share file, and established dissemination practices that were designed to meet the needs of most researchers.<sup>6</sup> The first article by Statistics Canada on the survey results came in September 1995,<sup>7</sup> the announcement of the release of the public use microdata data file came in November 1995,<sup>8</sup> and the first graduate thesis using NPHS was granted in the Spring of 1996.<sup>9</sup>

The NPHS had not been in the field long before new data gaps were identified and addressed through a new biennial cross-sectional survey, the Canadian Community Health Survey (CCHS). As the National Population Health Survey ceased to produce public use microdata files and use of the data was restricted to those who met the Statistics Canada's Research Data Centres criteria, the CCHS provided person-oriented health data at the sub-provincial level (health districts) and accommodated the need for periodic inclusion of special topics and special populations, such as mental health component of CCHS Cycle 1.2 for both the general population and the Canadian Forces.

**Methodology**

At the outset of this research project, the methodology used to discover evidence of NPHS use was that of standard literature reviews. Established peer-reviewed databases

were systematically searched to find proof of data use and knowledge transfer related to the NPHS. There was compelling evidence from this research that the data file was being used extensively, just as the survey planners had intended. Since the inception of the survey, there have been over 60 theses and dissertations written, over 400 articles published in 147 journals around the world, and in subject areas as varied as veterinary science, kinesiology, and cardio-thoracic research. Workshops and conference presentations abound, as do examples of the use of the data files in teaching health research courses.

The research also revealed the use of information related to the early stages of the survey in analytical studies. The research methodology was accordingly expanded to fit the framework of the survey life cycle (Diagram 1). While there are many representations of the life cycle of a survey, this framework contains those broad stages that are familiar to all project managers: the data gap analysis, the planning and administration of the survey, the release of the survey data and results, and finally, the evaluation of the survey in terms of its future. Search techniques and sources were revised to uncover products generated from all stages of the life cycle.<sup>10</sup>

**Findings**

The research revealed much more than simply articles

<b>NPHS Codebook (Cycle 2) Table of Contents</b>	
Background Objectives Survey Content Sample Design Data Collection Data Processing Data Quality Guidelines for Tabulation, Analysis & Release	Approximate Sampling Variability Tables Weighting File Usage Questionnaire Record Layout (General and Health) Data Dictionary Derived and Grouped Variables CV Tables List
<b>Diagram 2</b>	

analyzing the data files for the NPHS. Statistics Canada employees produced approximately 30 articles or reports on various aspects of the survey, dealing with topics as disparate as sample retention, consideration of weighting techniques, evaluation of statistical packages for NPHS analysis, the creation of dummy files, and the protection of confidentiality. There was no argument that the survey was successful in the public aspect of knowledge transfer; the database now totals some 800 items relating to the NPHS. The follow-up question was why the survey was so successful in meeting its initial goals. Discussions with the NPHS team yielded more information on the extent of internal documentation, such as training manuals, and interviewer feedback reports that help explain this extraordinary success.

**a) Survey Guides**

The detailed user guide accompanying the public use microdata file reflects the information that Statistics Canada considered necessary for data analysis (Diagram 2). It is notably more complete than most documentation disseminated with data files. The guide for NPHS Cycle 2 runs to 1060 pages (though admittedly, the inclusion of the Ontario Health Survey questions added bulk to the codebook). Nevertheless, does it have all the information that is going to help researchers analyze the survey properly? Are the traditional documentation and data files enough to ensure appropriate use of the survey data in future years?

Though the user manual for this survey is extremely rich, I would argue that it is not enough simply to capture the user guide for preservation. At a micro level, information on the question content and interview training is invaluable for focused research; at a macro level, life-cycle information provides context for understanding (and guiding) critical issues throughout the life of the survey.

**b) Background Reports/Training Materials/Admin-**

**istrative Documents**

There is an abundance of material produced across the life course of a survey that can help researchers understand the data. In preparing for the NPHS, for example, there were reports from the expert groups on selected areas of the questionnaire, developmental studies on new areas of health, questionnaire focus group reports, Treasury Board documents, studies on sample selection and methods to be used for estimates, coding concordances between survey cycles, analytical exercises for researchers using different software, and the results of post-field feedback surveys for interviewers.

Statistics Canada does pay special attention to training interviewers working on critical surveys, such as the Census of Population or surveys requesting sensitive information. The interviewer training materials for Cycle 1.2 of the Canadian Community Health Survey were particularly thorough because the survey's content focused on the state of mental health on a national level—a true challenge of the skills and training of Statistics Canada interviewers. Protocols in the manual included the handling of difficult respondents, referrals of distressed individuals to a list of resources, and contact information for mental health professionals for the use of interviewers to allow them to decompress after difficult sessions.

In the example presented in Figure 1, the CCHS Cycle 1.2 training manual provides information about handling sensitive questions. Is this information important to a researcher? Depending on whether this information relates to a question in her or his research area, it could be. One of the significant problems in investigating physical or substance abuse is presenting questions in a way that captures accurate replies without creating a biased social response.

Why is this useful information? Although there are many strategies for handling sensitive questions, it is extremely

DEP\_Q26EE1A During the last 12 months, did EXPERIENCE C happen to you?

R.: Well, yes, there was one night, we were drinking and I just...

Q. At this point the respondent stops talking and you see a tear come to her eyes. What do you do?

Exercise-Minimize Non-Response Tactics: Sad/Upset Respondent:

Stop for a minute. Be responsive to the respondent in a supportive way; give the respondent a chance to collect themselves, and help them get on track with the interview.

1. "That must have been very upsetting".....offer to take a break from the interview.
2. Ask if the respondent is able to continue with the interview.
3. If not, offer to continue the interview at a later date.

**Figure 1 :Training Exercise CCHS**

valuable when interpreting the survey results to know the specific instructions given to the interviewers. For instance, if the survey was interrupted and the interview continues with a subsequent denial of a previously positive answer, how was this handled?

In another example from the NPHS manual, interviewers are told that feedback on a question of social support indicated that the questions were problematic for members of certain cultures. In such cases, interviewers were to put "No response" rather than receiving and recording a negative answer.<sup>13</sup> Why is this potentially useful information? Immigrant studies attempt to identify factors that lead to a sense of belonging and social support. If particular cultures regard social support networks as not being applicable in their lives, it will help our understanding of the functioning of these cultures within and alongside the overall Canadian cultural norms. If the text or context of the question has prompted a consistent non-response pattern for a cultural group, that is equally important information.

### c) **Methodological Testing and Documentation**

On a broader level, the background of the NPHS is a compelling one, and understanding the role of the NPHS within the revamped health information infrastructure highlights the importance of issues under debate today. Two critical problems at this point, for example, are the attrition in the sample and the issues of privacy and data quality concerned with file linkage of the NPHS to administrative data.

Initial funding was provided for a sample size of 12,767

individuals, to be tracked over 20 years. A critical problem with the NPHS as it heads into the mid-point of the life of the survey is attrition within the original sample. The research on methods to retain the respondents of the original sample and the decisions on how best to make the survey sample robust are vital to the continued viability of this survey. The research community, for its part, needs to understand the efforts invested toward this end and support those efforts.

I would argue that researchers benefit in their analysis by knowing the context of the original sample selection and by understanding the dynamics behind the retention of the survey respondents. In this way, they can frame their questions appropriately and interpret analysis accurately. This knowledge can also contribute to research on sample selection for large longitudinal surveys. Much of this information is not in the current documentation but has been discussed in internal documents produced in conjunction with this survey.<sup>14</sup>

In supporting the development of the NPHS in 1991-1992, the Chief Statistician cited linkage as a key part of a robust health information system. It was assumed, for example, that certain causes of death in the mortality database might be traced back to NPHS respondents, and causal relationships on risk factors could be investigated more thoroughly. In the information leading up to the survey, it was argued that a direct product of linkage would be the improvement in the key determinants of health and a more effective use of health resources. Linkage is not a simple issue, though, particularly in a country that has responsibility for health split between the federal and ten

provincial governments—and three territories. The barriers and potential solutions to linkage issues inevitably escape the survey user guide, but are recorded elsewhere.

Soliciting information on how the survey producer and researchers want to see the data analyzed may help to determine the kinds of information to preserve throughout the survey life cycle. If, as the Task Force on Health Information asserted, analysis is a process to increase human perception of the significance of data, it is easy to see the degree of richness these additional documents offer to data analysis and interpretation. Ancillary documentation can help frame an issue so that the analysis is more meaningful. Consider, for example, the value of knowing the following types of information:

- What solutions are there to data linkage issues? Is there a response at the researcher or policy level that can alleviate technical and legal obstacles?
- Why did the wording of a question change over time? Are the responses still comparable or did the improvement in the clarity of the question change the distribution for this question?
- Why this particular question? Why not another wording that would give greater precision in the response?
- Why are respondent numbers declining and how do we maintain the reliability of the results?
- Why this content? Why not broader or more specific content?
- What scales were adopted or modified for use in this survey? Why?
- What can the interviewers tell us about the receptivity to more health questions, of geographical or cultural problems with data gathering at an individual or community level?
- How were difficult issues resolved? What was considered—and what was rejected?
  
- What is the relationship between this survey and other surveys of similar subject matter? Does this question, or order of questions, intentionally replicate the wording in other surveys?
- How has the sample changed (over surveys and over time) at the micro level? Who are those who have ceased to participate? Do they have lower incomes? Those in better health? Frequent movers?
- What should I know to analyze this survey appropriately?

In discussions with the project teams of the NPHS and CCHS, it is clear that they invested a great deal of talent

and energy into the design and administration of these surveys. It is not clear, however, that they always appreciate the value of the supporting documentation. I would argue that the documentation associated with the survey life cycle is too critical to be separated from the data itself. But does it all have to be included in the User Guide or relevant equivalent (such as DDI encoding)? Not necessarily, but while some documents are inappropriate for public release because of confidentiality and disclosure risks, researchers should know, at a minimum, what has been produced.

### **Intellectual and Physical Control**

At this point, none of the user guides produced by Statistics Canada include a complete listing of the documentation associated with a survey's development. The Integrated Metadata Data Base (IMDB) has become more inclusive with regard to documentation produced as part of a survey, but it does not come close to recording all the studies, reports, and evaluations produced in the life cycle of the survey.<sup>15</sup> With continuous surveys, the personnel associated with the surveys will move on. At best, their expertise and knowledge can only be partially transferred to their successors. Ironically, those who least appreciate the usefulness of the evolutionary work involved in a project are those who are responsible for the creation of the survey and for the ancillary documents involved in its creation, production, review, and revision in the first place.

Another key question is whether all of the documentation has to be physically stored or made accessible with the survey data. I would argue that the physical storage is not as critical as the intellectual linkage between the information produced in conjunction with a survey and the final version of the documentation and data file. This is a critical concept to grasp. Not all researchers need all of this information—in fact, it can be a disservice to those just beginning to work with a survey. However, the person investigating asthma will want to know that there is conceptual background for the content, that there is documentation on the variations in questions on asthma or asthma medications from cycle to cycle, and that there is a very good explanation as to why the question on pet ownership was omitted after the first cycle.

These examples refer to documentation directly related to the production of a survey, but the existence of other information is of equal value. Certainly the detailed survey methodology must be preserved even though the survey creators cannot release much of this information because of confidentiality constraints. Similarly, the documents prepared for the public release microdata committee may have similar confidentiality concerns. The need identified at this juncture is simply to know about the existence of this information by recording and linking it intellectually to the survey. Failure to understand that intellectual control is the highest priority puts information at risk of being suppressed, lost, or discarded as ephemeral, unimportant, or

dangerous.

Much like a publicly available report released in conjunction with a Royal Commission, the enduring studies provide lists of evidence and ensure that the background documents are all captured and archived. While the report itself will be preserved as part of the traditional function of the national library, the archives provide the security and archival techniques to preserve the audio, textual, and electronic forms of a commission's work.

### Conclusion

My research on the use of the NPHS and CCHS has indicated that the intensive planning involved in the survey has produced an extremely rich data file that has endured through the first half of the project. It has also unearthed a wealth of information that points to reasons why the NPHS, in particular, has been so successful in ensuring that the data are used. It is axiomatic that data are only valuable when analyzed and interpreted. Analysis and critical assessment of the data are the true mark of success of a survey.

As front line data curators, librarians, and archivists, our community champions access as a fundamental principle of democratic societies, a principle that, in practice, encourages ground-breaking investigation and research. It is also a cost-effective and sustainable way to produce an understanding of our society. Statistics Canada has neither the funding nor the mandate to conduct research, particularly in controversial areas.

As IASSIST members push forward in establishing protocols in the description of survey data for discovery and retrieval through DDI, it is incumbent upon us to think about the parameters of the information to be included in DDI information. Though it may not be imperative to capture and encode all the information generated through the life cycle of the survey, it is impossible to make an informed decision about inclusion and exclusion without knowledge of what was produced.

On their part, the research community has a mandate to delve into societal problems, and also the infrastructure to support long-term projects, including the type that will eventually produce the results needed to reshape a massive challenge like health care in Canada. Providing the resources to do this research is necessary; but the knowledge of the documents that exist to help explain and contextualize the variables and the data file is nothing short of critical to the success of this joint endeavour.

While data professionals may understand this better than the survey producers, it is worth asking whether archiving the documentation through the life cycle of a survey matters to the rest of the world. I would contend that it does. In a recent Senate Committee hearing, several

senators questioned Statistics Canada on the interview techniques used, and why there were different numbers used in the testimony of others appearing before the Committee.

Senator Lavoie-Roux: *My questions have already been asked more or less by my colleagues because I was wondering how the data were collected. It is easy to establish how many children complete their primary, secondary or university schooling because diplomas or certificates are granted. To me it was really an important issue to know to what extent your statistics could be trusted—I don't say that in a negative way—since the methods used for data collection are fairly weak. ...*

With the methods that you use, do you believe that we can trust the data? Are you sure that it is accurate?"<sup>16</sup>

It would be reassuring to say that yes, we know that we can trust the data. We have a complete record of evolution of that question, from concept to analysis.

### References

- Beaudoin, Gabrielle. "National Population Health Survey, Cycle 6: Respondent Relations." Workshop on Respondent Relations, Statistics Canada. Ottawa. February 23, 2004.
- Beland, Y. (2002). "Canadian Community Health Survey: Methodological Overview." *Health Reports*, 13, 3 (2002): 9-14.
- Belanger, Alain, Berthelot, J.M., Martel, L. "Canadian National Population Health Survey and the Calculation of Canadian Healthy Life Expectancy Indicators for the Next Decade." International Network of Health Expectancy, REVES 11 Conference. London. 1999.
- Bernier, Julie et al. "Comparison of the Health Utility Index Mark III (HUI III) and the Self-Assessed Health Status in the Canadian Population." International Society for Quality of Life Research (ISOQOL) Annual Conference. Amsterdam. 2001.
- Bernier, Julie et al. "Handling Missing Data." Statistical Society of Canada Annual Meeting. Toronto. 2002.
- Canada. Parliament. Senate. Standing Committee on Social Affairs, Science and Technology. *Proceedings: Evidence*. 36th Parliament, First Session: Issue 11, 13 May 1998.
- Canada. Task Force on Health Information. *Health Information for Canada, 1991: Report of the National Task Force on Health Information*. Ottawa: Health Canada, 1991.
- Canadian Institute for Health Information. *Health Information Roadmap: Beginning the Journey*. Ottawa: The

- Institute, 1999.
- Canadian Institute for Health Information. National Consensus Conference on Population Health Indicators Final Report. Ottawa: The Institute, 1999.
- Catlin, Gary. "Shaping a Vision for Health Statistics." National Committee on Vital and Health Statistics 50th Anniversary Symposium. Washington. 2000. <<http://ncvhs.hhs.gov/50thcatlin.pdf>> (25 June 2005).
- Catlin, G., & Will, P. "The National Population Health Survey: Highlights of Initial Developments." *Health Reports*, 4, 3 (1992): 313-329.
- Conference of Deputy Ministers of Health (Canada). Federal/Provincial/Territorial Advisory Committee on Population Health. *Strategies for Population Health: Investing in the Health of Canadians*. [Ottawa]: The Committee, 1994.
- Dale, A., Arber, S. & Procter, M. *Doing Secondary Analysis*. London: Unwin Hyman, 1988.
- Fobes P. & Geran L. "Cycle 2 and Beyond: Preparing and Storing Longitudinal Data of the National Population Health Survey." *Statistics Canada Symposium 98. Longitudinal Analysis for Complex Surveys*. Ottawa. 1998
- Gentleman, J., & Tomiak, M. "The Consistency of Various High Blood Pressure Indicators Based on Questionnaire and Physical Measures Data from the Canada Health Survey." *Health Reports*, 4,3 (1992): 293-312.
- Helmer, D. "Shooting From the Hip or Target Practice? A Comparison of Conventional and Fugitive Search Results." *Medical Library Association/Canadian Health Library Association Annual Conference*. Vancouver. May 2000.
- Helmer, D., et al. *Bibliography on Systematic Reviews: Locating, Understanding and Using the Evidence*. BCOHTA Working Document, no. 1. Vancouver: Office of Health Technology Assessment, 1999.
- Helmer, D., Savoie, I., & Green, C. "Evidence-Based Practice: Extending the Search to Find Material for the Systematic Review. *Bulletin of the Medical Library Association*, 89, 4 (2001): 346-352.
- Helmer, D., Savoie, I., & Green, C. "How Do Various Fugitive Literature Searching Methods Impact the Comprehensiveness of Literature Uncovered for Systematic Review?" *International Conference on Grey Literature*. Washington, DC. 1999.
- Mantel, H.J. & Nadon, S. "Dummy file creation for the remote access program of the National Population Health Survey." In *Statistics Canada. Survey Methods Section. Proceedings*. Ottawa: Statistics Canada, 1999.
- Mathieu, Patrice. "Longitudinal Research Using the National Population Health Survey." *Situating Place in Health Research: Putting Theory into Practice*. Kingston. 2001.
- Norris, D. "Using Survey Data to Address Policy Issues Related to Health and Social Support." *Canadian Association on Gerontology Annual Scientific and Educational Meeting*. Halifax. 1998.
- Perez, Claudio E. *Using the Bootstrap Technique for NPHS Analysis*. Ottawa: Health Statistics Division, Statistics Canada, n.d.
- Roberts, Georgia, et al. "Bridging the Gap Between the Theory and Practice of Analysis of Data from Complex Surveys: Some Statistics Canada Experiences." *Federal Committee on Statistical Methodology Research Conference*. Washington. 1999.
- Savoie, I., & Helmer, D. "Extended Systematic vs. Conventional Search Methods: Weighing the Quality of the Literature Retrieved." *International Society for Technology Assessment in Health Care Annual Meeting*, Philadelphia, PA. 2001.
- Savoie, I., Helmer, D., Green, C.J., & Kazanjian, A. "Improving the Efficiency of the Fugitive Search: Effectiveness of Various Methods. *International Society for Technology Assessment in Health Care Annual Meeting*. The Hague. 2000.
- Scherer, R.W., & Langenberg, P. "Full Publication of Results Initially Presented in Abstracts." *Cochrane Methodology Review*, 2 (2001).
- Shields, Margot. "Proxy Reporting in the National Population Health Survey." *Health Reports*, 12, 1 (2001):21-40 (English), 23-44 (French).
- Singh, M.P. et al. *National Population Health Survey: Design and Issues*." In *American Statistical Association. Section on Survey Research Methods. Proceedings*, 1994. Alexandria: American Statistical Association, 1994. [http://www.amstat.org/sections/srms/Proceedings/papers/1994\\_138.pdf](http://www.amstat.org/sections/srms/Proceedings/papers/1994_138.pdf)> (25 June 2005).
- Statistics Canada. *Canadian Community Health Survey, Cycle 1.2: Mental Health and Well-being, Project Code 6502-0 Training Guide*. Ottawa, Statistics Canada, [n.d.].
- Statistics Canada. *Information about the National Population Health Survey*. (82F0068XIE). Ottawa: Statistics Canada, 1999. <http://www.statcan.ca:8096/bsolc/>

english/bsolc?catno=82F0068X (Accessed 1 July 2005).

Statistics Canada. *National Population Health Survey 1996-97 Household Component: User's Guide for the Public Use Microdata Files*. Ottawa: Statistics Canada, 1998.

Statistics Canada. *National Population Health Survey Answers to the Exercise to Minimize Non-Response, Cycle 6, Project Code 0104-0*. [Ottawa: Statistics Canada, n.d.]

Statistics Canada. *National Population Health Survey Debriefing Questionnaire 6, Cycle 6 – Quarter 4 – 0104-0*. [Ottawa: Statistics Canada, n.d.]

Statistics Canada. *National Population Health Survey Interviewer's Manual, Cycle 6, Project Code 0104-0*. [Ottawa: Statistics Canada, n.d.]

Statistics Canada. *National Population Health Survey Self-Study, Cycle 6, Project Code 0104-0*. [Ottawa: Statistics Canada, n.d.]

Statistics Canada. *National Population Health Survey Training Guide, Cycle 6, Project Code 0104-0*. [Ottawa: Statistics Canada, n.d.]

Statistics Canada. *Cycle 4 (2000-2001) Household Component Longitudinal Documentation*. Ottawa: Health Statistics Division, Statistics Canada [unpublished], May 2002.

Stephens, T. Measuring the Health of Canadians: an Agenda for Developing Health Surveys. *Health Reports*, 3, 2 (1991): 137-145.

Stukel, D.M., Mohl, C. & Tambay, J-L. "Weighting for Cycle Two of Statistics Canada's National Population Health Survey." Statistical Society of Canada. *Proceedings of the Survey Methods Section*. Ottawa: Statistical Society of Canada, 1997.

Swain, L., Caitlin, G., & Beaudet, M.P. The National Population Health Survey – Its Longitudinal Nature. *Health Reports*, 10,4 (1999): 69-82.

Tambay, J-L., & Catlin, G. "Sample Design of the National Population Health Survey." *Health Reports*, 7, 1(1995): 29-38.

Tambay, J-L. et al. "Treatment of Nonresponse in Cycle Two of the National Population Health Survey." *Survey Methodology*. 24 (1998): 147-156 (English), 159-169 (French).

Tolusso, S. & Brisebois, F. *NPHS Data Quality: Exploring Non-sampling Errors*, Methodology Branch Working Paper,

HSMD-2003-004E. Ottawa: Statistics Canada, 2003.

Umphrey, G., Kendall, O., & MacNeill, I.B. "Assessing the Surveillance Capability of Canada's National Health Surveys." *Chronic Diseases in Canada*, 22, 2 (2001): 50-56.

Wolfson, Michael C. "Toward a System of Health Statistics" *Daedalus*. 123 (Fall 1994): 181-195.

Yeo, D. "After the First Steps: the Evolution of a Longitudinal Survey." Workshop on Longitudinal Research in Social Science—A Canadian Focus. London, ON. 1999.

Yeo, D., Mantel, H. & Liu, T.P. "Bootstrap variance estimation for the National Population Health Survey." In Proceedings of the American Statistical Association. Survey Research Methods Section. Alexandria, VA: The Association, 2000.

#### Footnotes

<sup>1</sup> Elizabeth Hamilton is Head, Government Documents, Data and Maps Department at University of New Brunswick. Contact: hamilton@unb.ca

The paper is a revision of material presented at the IASSIST 2005 conference in Edinburgh.

<sup>2</sup> The context of a survey has been recognized by authors such as Dale, Arber and Procter (1988) as important in secondary analysis, but has received little attention within the context of identification, documentation and preservation of master files and public use microdata files.

<sup>3</sup> For further information on the background of the NPHS, see: D. Yeo, After the first steps : the evolution of the National Population Health Survey, *Canadian Studies in Population*, 28; 2 (2001):.377-390.

<sup>4</sup> "The lack of standard data definitions, minimum data sets, standard edit rules, as well as the quality control and security procedures is completely unacceptable." Extract from the report of the Project Team on Comparability, in Canada. Task Force on Health Information. Health Information for Canada, 1991: *Report of the National Task Force on Health Information* (Ottawa: Health Canada, 1991), 9.

<sup>5</sup> The research process had begun long before the approval of funding, of course. For example, blood pressure indicators had been a subject of study in the early 1990s, and there had been a Federal-Provincial-Territorial committee on mental health before NPHS was born.

<sup>6</sup> Two of the programs that evolved from the desire to promote timely access to data were the highly successful Data Liberation Initiative (DLI) launched in 1996, and



the Research Data Centres (RDCs). NPHS project managers contributed products through all Statistics Canada's dissemination channels, including tables in *Health Indicators*, published articles (in both free and priced publications), public use microdata files through DLI, master data files through the RDCs, and remote job submission and custom tabulations for individualized data tabulations.

<sup>7</sup> "Health of Canadians, 1994" in *The Daily*, 22 September 1995

<sup>8</sup> *The Daily*, 21 November 1995

<sup>9</sup> Paula Fletcher, *Falls among the elderly: risk factors and prevention strategies*. PhD, University of Waterloo, 1996

<sup>10</sup> See, for example, D. Helmer, I. Savoie, & C. Green, "Evidence-Based Practice: Extending the Search to Find Material for the Systematic Review." *Bulletin of the Medical Library Association*, 89, 4 (2001): 346-352; and D. Helmer, "Shooting From the Hip or Target Practice? A Comparison of Conventional and Fugitive Search Results." Medical Library Association/Canadian Health Library Association Annual Conference. Vancouver. May 2000.

<sup>11</sup> Statistics Canada. *National Population Health Survey 1996-97 Household Component: User's Guide for the Public Use Microdata Files*, (Ottawa: Statistics Canada, 1998), 3-5..

<sup>12</sup> Statistics Canada. *Canadian Community Health Survey, Cycle 1.2: Mental Health and Well-being, Project Code 6502-0 Training Guide* (Ottawa, Statistics Canada, [n.d.]) 183-184.

<sup>13</sup> Statistics Canada. *National Population Health Survey: Cycle 6: Project Code 0104-0: Training Guide* (Ottawa: Statistics Canada, [n.d.]), 124.

<sup>14</sup> Beaudoin, Gabrielle, "National Population Health Survey: Cycle 6 Respondent Relations (Ottawa: Statistics Canada, 2004).

<sup>15</sup> The entry for the CCHS includes the information that there were studies evaluating CCHS results in relation to various other surveys, and a study that compared the profile of respondents agreeing to share their data with other departments was conducted. <http://www.statcan.ca/english/sdds/3226.htm>

<sup>16</sup> Canada. Parliament. Senate. Standing Committee on Social Affairs, Science and Technology. *Proceedings: evidence*. Issue 11, 13 May 1998.