Barriers to Accessing & Analyzing Health Information in Canada

November 2002

Dr. George Kephart
Dalhousie University

CPHI SPC
Canadian Population Health Initiative | Initiative sur la santé de la population canadienne

Canadian Institute for Health Information
Institut canadien d'information sur la santé

Barriers to Accessing & Analyzing Health Information in Canada

Dr. George Kephart
Dalhousie University

November 2002

Canadian Institute for Health Information
377 Dalhousie Street
Suite 200
Ottawa, Ontario
K1N 9N8

Telephone: (613) 241-7860
Fax: (613) 241-8120
www.cihi.ca

ISBN 1-55392-117-8 (PDF)

Cette publication est disponible en français sous le titre : « Les obstracles à l'accès et à l'analyse de l'information sur la santé au Canada » ISBN 1-55392-118-6 (PDF)

Barriers to Accessing & Analyzing Health Information in Canada

Table of Contents

Acknowledgements

Research Contact

Dr. George Kephart
Department of Community Health and Epidemiology
Faculty of Medicine
Dalhousie University
5849 University Avenue
Halifax, Nova Scotia, B3H 4H7
Email: George.Kephart@Dal.ca

Principal Investigators

Dr. George Kephart, Dalhousie University
Dr. Rob Reid, University of British Columbia
Dr. Nazeem Muhajarine, University of Saskatchewan
Dr. Leslie Roos, Manitoba Centre for Health Policy and Evaluation
Dr. Michael Wolfson, Statistics Canada
Dr. Jack Williams, Institute for Clinical Evaluative Sciences
Dr. Doug Manuel, Institute for Clinical Evaluative Sciences

This project was supported by a financial contribution from the Health Transition Fund of Health Canada. It was also supported by the Canadian Population Health Initiative, and by Statistics Canada. The views expressed herein do not necessarily represent the official policy of federal, provincial or territorial governments, nor do they represent the official policy of the Canadian Population Health Initiative/Canadian Institute for Health Information.

Executive Summary

Purpose

Given the diversity of approaches to health care and efforts to reduce health care expenditures across the country, concerns about equity in access to health care are pertinent on the Canadian health research agenda. This project combined federal and provincial resources to develop the data, methodological tools and expertise necessary to study and monitor socio-economic differences in access to health care. The goal was to assemble a federal/provincial organizational structure and an integrated, cost-effective federal/provincial data resource that could be used for future research.

Previous projects integrating federal and provincial health data were smaller in scale and were conducted by national organizations such as the Canadian Institute for Health Information (CIHI), Statistics Canada and Health Canada. This project, on the other hand, while including CIHI and Statistics Canada as partners, was conducted by a collaborative network of university-based researchers in five provinces: British Columbia, Saskatchewan, Manitoba, Ontario and Nova Scotia. It was hoped that this innovative approach would help to navigate inter-provincial and federal differences in policies governing data access and sharing and to overcome inter-provincial barriers to analysis arising from differences in the concepts and definitions used in provincial databases.

The research plan for the Project involved creating a data set linking a Statistics Canada survey, the National Population Health Survey (NPHS) with provincial hospital discharge and doctor administrative records for all those individuals surveyed in the NPHS who had given their prior consent for such linkage, in the five provinces with the participating researchers. Each researcher would manage the data assembly, permissions and approvals processes in their own provinces. Once the province-by-province data assembly was completed, all five files would be sent to Statistics Canada, where the files would be pooled. All subsequent access to the linked and pooled data would only be allowed on Statistics Canada's premises by "deemed" Statistics Canada employees, hence under the full protection of the *Statistics Act*. In other words, all the provincial researchers working directly with the data would be sworn in as Statistics Canada deemed employees so the analysis and research could proceed.

This report documents the process of assembling the data for this project, describes the logistical and organizational barriers to combining federal and provincial data resources and expertise and offers recommendations on how to overcome these barriers.

Target Audience

The target audience for this report includes the following federal and provincial organizations that need to be informed about the challenges and barriers to accessing health data for the purpose of conducting population health and health services related research: Federal/Provincial/Territorial Working Group on the Protection of Personal Health Information; Health Canada; Statistics Canada; Canadian Institute for Health Information; Provincial Ministries of Health; Privacy Commissioners; Federal Provincial/Territorial committees reporting to the Conference of Deputy Ministers of Health.

Methods and Activities

The project linked data from the 1994 and 1996 waves of the National Population Health Survey (NPHS) with provincial physician and hospital utilization data for the years 1992–1998. Respondents to the NPHS had explicitly provided their consent for linkage of their survey data to administrative data. Despite this consent, the diversity of the approvals required in each province to proceed with the project while ensuring the privacy of health information, and the lack of procedures and policies to facilitate data sharing between provinces, introduced long delays into the project. Problems in comparability between provincial health data were also a major barrier that affected the usability of physician claims data, and to a lesser degree of hospital discharge data.

Organizing data access through Statistics Canada provided a useful mechanism for sharing data between federal and provincial governments. At the same time, Statistics Canada's *On Premises Access Microdata Research Contract* which researchers were required to sign introduced substantial barriers and delays because of requirements that were inconsistent with university policies, and contracted obligation to Health Canada. These problems, which may hamper the ability of Statistics Canada to participate in research partnerships, were addressed through changes to the contract.

Main Findings and Benefits

The linked data for this project combine the best information available on the demographic, socioeconomic and health characteristics of Canadians with the most detailed information available on health care utilization in administrative health care databases. By combining information from national surveys with provincial health information through an effective collaborative framework, this project created research opportunities not previously available in Canada, and not generally available internationally, enabling researchers to:

- Study the effects of health determinants on health services utilization and health outcomes.
- Study the effects of health events and health care utilization on factors such as employment, income and psychosocial characteristics.
- Make inter-provincial comparisons of health status, access to health services, health outcomes, etc.

Recommendations and Policy Implications

The organizational framework used in this project should be employed in future national research efforts to effectively link federal and provincial data and expertise. The promotion of successful federal/provincial health services and population health research efforts requires federal/provincial/territorial research partnerships that:

- Draw expertise from the federal government, the provinces and academic research institutions.
- Link federal and provincial data resources.

However, these partnerships need to be augmented with federal/provincial/territorial organizational support to:

- Set priorities to help ensure support for the project at multiple levels of government.
- Harmonize review procedures, particularly around privacy and confidentiality.
- Develop more effective and streamlined policies and contractual arrangements governing the sharing of sensitive data between Statistics Canada and provincial governments.

Recommendations for Statistics Canada

Statistics Canada should explore ways to make their contracts with outside researchers more flexible by broadening their definition of the "product" required. For example, in cases where Statistics Canada is an official partner in research initiatives, perhaps the "product" could be defined as Statistics Canada's deliverable to the partnership (e.g. a report released by another federal agency, such as Health Canada, with appropriate acknowledgement of Statistics Canada's contribution).

Recommendations for CIHI

The development of data standards and of ways to bridge coding differences should continue to be a high priority for CIHI. In particular, the research demonstrates the importance of ongoing CIHI initiatives to establish comparable data collection standards and protocols for different types of claims data and health information. CIHI and the federal government should also place a high priority on the broad dissemination of technologies that facilitate inter-provincial and longitudinal health research, such as the National Grouping System and CMG™ software.

Recommendations for the Provinces

Provincial governments need to develop harmonized policies and procedures for accessing and sharing data.

Introduction

On a broad scale, the primary goal of this project was to assemble a federal/provincial organizational structure, and an integrated federal/provincial data resource, that could be used for research that extends beyond the current project. The general objective was to build a cost-effective data resource and a collaborative framework of health analysts and researchers who could break down inter-provincial barriers to analysis arising from differing concepts and definitions in their provincial data bases, as well as from inter-provincial and federal differences in policies governing data access and sharing. This broad objective was accomplished by means of face-to-face meetings, extensive electronic communications and innovative strategies for data sharing, security and access that helped to facilitate the sharing and combination of knowledge, expertise and data.

This report documents the process of assembling the data and the expertise necessary for this project. It focuses on the logistical and organizational barriers to accessing, linking and analyzing federal and provincial data resources and provides recommendations for addressing these barriers. The importance of the project extends beyond an examination of access to health care services. The lessons learned about the organizational challenges overcome and the data gathered provide an important resource for an array of health services and population health studies. The dissemination of empirical results is being done through Statistics Canada and the Canadian Population Health Initiative.

Background: Original Research Objectives

Preventive health care is one area within the health care sector where clear opportunities for addressing socio-economic inequities exist. This project sought to examine the extent to which lower socio-economic groups are less likely to receive preventative health services and to explain the reasons for this inequity. Such knowledge is critical in the design of programs and policies to improve the quality of preventative health care for those groups, which are currently at higher risk of disability and death from preventable diseases.

The first research objective of the project was to:

Develop and examine a set of indicators of equity in access to health services that could be used in the evaluation of the core services of Canada's health care system.

More specifically, the goals were to:

a. Assess equity in access to health services by socio-economic status (SES) (i.e. SES differences in utilization net of health status) for:
    - provinces
    - urban and rural areas
    - age and sex

b. Explore financial and non-financial barriers that may account for any observed inequities in access to health services. Factors to be examined include:

- living arrangements
- stress (stress associated with finances, work and children)
- labour force participation
- psychosocial factors (e.g. sense of coherence, mastery and self-esteem)
- social support

The second research objective was to:

Develop and examine a set of indicators of SES differences in the use of preventative health services in order to:

a. Evaluate socio-economic differences in the use of the following preventive services:

- cervical cancer screening
- mammography
- clinical breast exam
- rectal exam
- flu shot
- diagnosis and treatment of high blood pressure

b. Assess socio-economic differences in the use of physician services by:

- age and sex of patient
- general versus specialist services
- urgent (emergency, evening and weekend) versus routine office visits
- primary reason for physician visits (based on primary diagnosis)

c. Explore the degree to which socio-economic differences in the use of preventive health practices are associated with:

- differences in the frequency and type of physician services received
- differences in the demographic and practice characteristics of physicians
- patient factors which may impose non-financial barriers to access

Background: Project Activities

Overview

The research plan for the Project involved creating a data set linking a Statistics Canada survey, the National Population Health Survey (NPHS) with provincial hospital discharge and doctor administrative records for all those individuals surveyed in the NPHS who had given their prior consent for such linkage, in the five provinces with the participating researchers. Each researcher would manage the data assembly, permissions and approvals processes in their own provinces. Once the province-by-province data assembly was completed, all five files would be sent to Statistics Canada, where the files would be pooled. All subsequent access to the linked and pooled data would only be allowed on Statistics Canada's premises by "deemed" Statistics Canada employees, hence under the

full protection of the *Statistics Act*. In other words, all the provincial researchers working directly with the data would be sworn in as Statistics Canada deemed employees so the analysis and research could proceed.

The Subjects for the study were the respondents to the 1994 and 1996 waves of the National Population Health Survey (NPHS) in the five provinces studied who provided consent to have their data linked with administrative data and to be shared by Statistics Canada with the provinces. The NPHS data include information on the subjects' health status and on their social and economic characteristics. The NPHS records which met these conditions in each of the five participating provinces were linked with provincial administrative data on their health services utilization. While Ontario data had been linked, the Ontario government did not authorize the release of the data to Statistics Canada, and thus it was not possible to include that province in analyses.

In order to proceed with the linkage, several interrelated tasks had to be completed. First, the specific study variables to be derived from the administrative data had to be identified. Second, a data flow and access protocol had to be developed to protect the confidentiality and privacy of study subjects. A description of the study, the protocol, and a detailed list of, and justification for, data requested for linkage were required in order to obtain Statistics Canada and provincial approvals. The approvals required to proceed with the study differed substantially between provinces, and included various combinations of ethics reviews, privacy/confidentiality reviews and ministerial approvals. The process of obtaining provincial approvals was slow, resulting in extensive delays to the project. In some cases, obtaining approvals necessitated changes in the protocol and data to be linked. In addition to provincial approvals, security clearances, contracts, policy changes and Statistics Canada approvals had to be obtained. In summary, the administrative requirements to simply proceed with the study were complex, variable and time consuming.

The study also had to address issues of data comparability in physician claims data and in hospital discharge data. Provinces differ in the coding systems they use for physician claims: procedures are coded differently, different coding schemes are used to classify physician specialties, there are differences in the amount and degree of diagnostic coding, and there are differences in the types of physician services that appear in the databases. Although CIHI has standardized the coding of hospital discharge data, there are nevertheless a number of significant coding differences between provinces in that type of data as well.

Data Sources

The project linked data from the 1994 and 1996 waves of the National Population Health Survey (NPHS) with provincial physician and hospital utilization data for the years 1992–1998.

National Population Health Survey (NPHS)

The purpose of the NPHS is to:

- Assess the health status of the Canadian population and its relationship to the determinants of health and to the utilization of health care services.

- Amass data on the social, economic, demographic, environmental, and occupational correlates of health.
- Follow a panel of individuals over time to obtain information on the dynamic processes of health and illness.
- Allow for the possibility of linking survey data to routinely collected administrative data (Swain, Catlin, and Beaudet, 1999).

The NPHS is representative of the Canadian population over 12 years old, with the exception of individuals living on Canadian Forces bases, Indian reserves, and individuals in the Yukon, Northwest Territories, and remote areas in Ontario and Quebec (Smith and Haynes, 1992), (Statistics Canada, 2001). The survey consists of a general component and a health component. The general questionnaire contains demographic, socioeconomic and limited health information on each person in the household, while the health questionnaire comprises in-depth health questions about one randomly selected individual in the household (Swain, Catlin, et al. 1999 ID: 531). The NPHS contains both cross sectional and longitudinal core data. The longitudinal data follows primary respondents from the 1994/1995 wave of the NPHS, 90.7% of whom were followed up in the 1996/1997 survey. Although a 1998 wave is now available for analysis, it was not used in this study. In both Manitoba and Ontario, supplemental cross-sectional samples were drawn in the 1996 wave of the NPHS. While the supplemental samples have been linked to administrative data, the analysis in this project was limited to the longitudinal sample.

From the NPHS data, this study obtained information that is not available in administrative databases on demographic characteristics, socioeconomic status (household income, education, occupation and employment status), non-financial barriers to health-care utilization and health status. A variety of variables from the NPHS were used as measures of opportunity costs, barriers and enabling factors that may affect access to health services. These include labour force participation (e.g. average hours worked per week and shift work), household structure (e.g. ages of children and single parent status), stress and psychosocial factors (sense of coherence, mastery, etc.). Health status, which along with age and sex is an important predictor of the need for health services, is being measured with a number of variables. These include the number and types of chronic health conditions, disability, self-reported health status, and the Health Utility Index.

Provincial Administrative Health Databases

The five provinces involved in the study have high quality health care administrative data that have been longitudinally linked and widely used for research purposes. Administrative health databases contain information on insured health services provided by provincial governments. Depending on the province, they may include information on physician services, prescription drug use, home care, long-term care and hospital stays. This study employed two sources of administrative data that are generally available in all provinces: fee-for-service physician claims data and hospital discharge data.

Physician payment information is essentially financial transaction data collected for purposes of physician payment. In all provinces, physicians are paid primarily on a fee-for-service basis according to a detailed list of fee codes which describe the kind of service performed and the cost of the service. Physician claims data typically contain information on the date and type of service performed, the amount paid, and basic diagnostic information. Claims also include patient and physician identifiers that permit linkage of the

data (across time, or to other databases). While these databases include the vast majority of physician services, they may or may not include services performed by physicians on alternate payment schemes, such as salaries. For the time period and provinces in this study, the impact of this problem is minimal except for the Hamilton-Wentworth area of Ontario, which could not be included in the study.

Hospital discharge data include information on all discharges for inpatient stays and day surgeries (although day surgeries are not defined uniformly across provinces). The data record the patient's health insurance number, postal code, admission and discharge dates, primary and secondary diagnoses, procedures, as well as information about where the patient was admitted from, and where he or she was discharged to. Thanks to the standards for the collection and coding of hospital discharge data established by CIHI, these data are relatively comparable across the study provinces. Nevertheless, a variety of comparability issues, which will be discussed in detail in the section on Standardization of Administrative Data, had to be addressed for this study.

From the provincial administrative databases, the project obtained detailed information on health services utilization. The self-reported data that the NPHS collected from respondents on their use of health services during the previous year are subject to recall error and do not provide information on the specific type of service used. Physician claims data, on the other hand, make it possible to describe physician utilization by specialty and type of service. Hospital discharge data provide detailed information on the type of stay (e.g. medical versus surgical), the associated diagnoses, and the length of stay. This allowed researchers to look at admission rates and stays for particular health conditions. The administrative data also allowed them to measure the utilization of health services prospectively (i.e. after the survey), and over a longer period of time (in this study, six years).

Barriers to Accessing Health Information: Approval Process and Contractual Agreements

Despite the fact that respondents to the NPHS had provided consent for linkage of their survey data to administrative health data, an extensive series of approvals were required to proceed with the project (see Appendix). The ethics reviews required in all provinces were conducted by University ethics committees, and guided by the Tri-Council Policy Statement, "Ethical Conduct for Research Involving Human Subjects". The ethics review weighs the potential benefits of the project against the potential harm to the study subjects. For this study, the primary risk was violation of the subjects' privacy and confidentiality, but that risk was minimal given the study protocol.

Unlike ethics reviews, the privacy and confidentiality reviews that were required in most provinces and in Statistics Canada to ensure the confidentiality and privacy of personal health information had no standard set of procedures, policies or criteria. They varied in each province based on internal policies, contracted responsibilities (for custodians of the data) and legislation. There was also considerable variability between the provinces in who conducted the privacy review:

- In Manitoba, the review was conducted by the Health Information Privacy Committee, which was established in accordance with *The Personal Health Information Privacy Act* of Manitoba. The Committee was unique among the provinces in having a legislated mandate and membership representing a variety of stakeholders (e.g. public representatives, health care providers and researchers).
- In Nova Scotia, the review was conducted by the Population Health Research Unit (PHRU) at Dalhousie University, which acts as the custodian of provincial health data for research purposes. It was conducted by PHRU management in accordance with their contractual obligations to their provincial government, and by the university ethics review board. The contract required that data be accessed and used in accordance with the Nova Scotia *Freedom of Information and Privacy Act*, 1993, c.5, s.1.
- In British Columbia, the review was conducted by the Data Access Subcommittee (DAS), established by the Centre for Health Care Policy and Research at the University of British Columbia, and by the British Columbia Ministry of Health. The DAS develops and administers policies and procedures to streamline data requests for linked data, while ensuring that the data use is in the public's interest, and that it complies with the British Columbia Freedom of Information and Privacy Act.
- In Saskatchewan, the processing and approval of data requests for research is reasonably streamlined. The Data Access Review Committee in Saskatchewan Health reviews all data requests to ensure compliance with the Department's data access and confidentiality policies. Furthermore, privacy legislation has been approved by the legislature, and is awaiting proclamation. Saskatchewan was by far the most restrictive province regarding the release of data for linkage. The province required, for example, that many data fields, especially those containing detailed information such as diagnoses and fee codes, be grouped into broader categories prior to release.
- In Ontario, the review was conducted internally by a privacy officer and the CEO of the Institute for Clinical and Evaluative Sciences. The Legal Branch of the Ministry of Health is also reviewing the request with respect to freedom of information and privacy. This review is still pending.

The time required for privacy and confidentiality reviews varied, but was generally less than a month following the submission of required information. However, since detailed information on the variables requested, the linkage protocol, and the methods for controlling data access and security were required with submissions, the preparations for submission were time consuming. In some cases, the review resulted in requests for additional information, or in changes to the study design and protocol. This necessitated coordination between the provinces because if one province, for example, required a change in protocol, other provinces may have not been able to make similar changes without re-submitting the changes for review. Preparing the submissions also required considerable advance planning at the outset of the project. Accurate identification of all the information needed was critical because "going back" to get additional information would require an additional cycle of reviews, data extractions and linkages. Also, the number of changes that could be made to the protocol was restricted.

In addition to ethics, privacy and confidentiality reviews, other provincial approvals and requests for assistance also introduced significant barriers and delays. In two provinces, Nova Scotia and Ontario, special approval was required to send data to Statistics Canada

from the provincial Ministry of Health. In Nova Scotia, special permission was obtained in writing from the Deputy Minister of Health with only small delays. In the case of Ontario, on the other hand, this requirement resulted in significant delays, and Ministry approval took nearly 1.5 years to obtain. The data release must also be approved by the Ministry's legal branch, and is still pending. In Manitoba, special approval was not required, but linkage had to be conducted within Manitoba Health subject to the Ministry's priorities and time constraints. As a result of linkage problems with the 1994 NPHS file, only respondents in the 1996 wave were linked.

A number of contractual arrangements and clearances were also required for the project to proceed. For example, Statistics Canada required all researchers and staff working directly with the data to obtain security clearances, and to be sworn in as "deemed employees" of Statistics Canada. Statistics Canada's *On-Premises Access Microdata Research Contract,* which they were required to sign, included provisions that were inconsistent with university polices on publication rights. The contract also required that the first dissemination of results be through a Statistics Canada publication. Submission of a report to the Health Transition Fund first would be a violation of the contract, and so contract revisions were required to allow for the submission of the report to the Health Transition Fund. The provinces also asked Statistics Canada to establish a policy imposing some restrictions on future use of the study data. In particular, they requested that any use of the data for other purposes would require additional provincial review, which resulted in further delays. In summary, a remarkably complex array of approvals, clearances, contractual arrangements and policy regulations had to be cleared before the project could proceed. As of November 2002, more than two years have been expended obtaining these approvals, and final approval from Ontario is still pending. Analysis proceeded using data from the other four provinces.

Standardization of Administrative Data

Before the administrative data could be linked, it was necessary to identify coding differences to ensure that the data submitted by the different provinces was comparable. Analysts in Nova Scotia and Manitoba, with assistance from the other provinces, developed data standards for the linked administrative data which specified the required and optional variables to be extracted and provided consistent coding of key variables.

Physician claims data imposed the most complex compatibility problems because there are substantial differences in the coding of diagnoses and activities across provinces. The choice of activities to code, the ways these activities are coded and the amount paid for each activity are to a large extent the result of negotiations between provincial medical societies and Departments of Health. In some provinces, for example Manitoba, British Columbia and (prior to 1997) Nova Scotia, largely independent coding systems were used. Ontario, Saskatchewan and (since 1997) Nova Scotia, on the other hand, have adopted a more standardized coding system, the CCP system.

Relatively early in the project, the study team recognized that resolving coding differences in physician claims data was a large undertaking that could not be accomplished within the given time and funding constraints. One option was to make use of the National Grouping System developed by CIHI which categorizes the fee codes for each province into a

standard set of comparable categories. Unfortunately, however, the National Grouping System was developed for the purpose of building the National Physician Database and is not set up as a product that can be easily disseminated to the provinces. Developing provincial capacity to implement the National Grouping System would have substantial benefits that extend beyond the current project. The team worked with CIHI to explore this possibility, and CIHI was prepared to provide substantial assistance to help with the implementation of the NGS, but it was concluded that this could not be accomplished in this phase of the project. The team will be seeking new funding sources to pursue this option in the future. For the purposes of this study, very basic measures of physician utilization were employed (number of general practitioner and specialist visits), and these may be subject to some bias from coding differences. A subset of physician services, those that are coded similarly in all study provinces, were examined. Physician services that are not coded, or that are coded differently in some provinces (e.g. radiology), were excluded.

Even though hospital discharge abstracts are considerably more standardized than physician claims data, there are nevertheless differences in the coding of data across time, and some remaining differences in coding between the provinces. The most important issues that affected the study were:

- Day surgeries are not coded consistently across provinces, and it was necessary to exclude some types of day visits to arrive at a common definition.
- CIHI uses information such as age and diagnostic information to classify patient stays into a set of "case mix groups," or CMG. Because the grouping procedure was continually updated, CMG were not coded consistently over time in most provinces. Software is available to group previous years of discharge data according to the most recent grouping procedure, but this process is expensive and was undertaken by only one of the five provinces—Ontario.
- Prior to 1995, Nova Scotia had its own coding system that was similar to the system developed by CIHI, but included a number of important differences. For example, it did not contain as much diagnostic information.

Data Extraction and Linkage Procedures

While some variations existed, the typical steps in the development of the linked data were as follows. Statistics Canada sent a file containing only health insurance numbers and NPHS identifiers to the provinces. If respondents provided consent to have their records linked, but did not provide their health insurance numbers, names and other identifying information were sent to the provinces. In some provinces (e.g. Nova Scotia), the identifying information was used to obtain missing insurance numbers. Using health card numbers, each province extracted and standardized the administrative data for NPHS respondents and sent the administrative data to Statistics Canada. The benefit of this approach was that NPHS data did not need to be combined with administrative data outside the Statistics Canada office. Statistics Canada then transmitted the data for all of the provinces, along with the NPHS data, to regional offices in each of the participating provinces. The regional offices have each established facilities, including workspace, computers and statistical software, enabling project staff to have supervised access to the data.

Provinces used different approaches, and applied various internal security procedures, to extract data for NPHS respondents. Data extraction in British Columbia was implemented by the study team at the Centre for Health Care Policy and Research at UBC using encrypted health card numbers. In Saskatchewan, the data extraction and linkage were conducted entirely by the provincial government. Saskatchewan was unique in that the investigators did not have direct access to their provincial health data. In all other provinces, the research teams worked in centers that serve as custodians of provincial data, and were able to manipulate and extract provincial data themselves before sending it to Statistics Canada. In Manitoba, linkage of NPHS health card numbers to a resident master file was conducted by Manitoba Health, but the actual data extraction was done by the members of the study team at the Manitoba Centre for Health Policy and Evaluation using encrypted health insurance numbers. Similarly, in Nova Scotia, an agent for the Nova Scotia Department of Health (Maritime Medical Care Inc.) looked up missing health card numbers, encrypted health card numbers using a unique algorithm, and sent the data to the Population Heath Research Unit at Dalhousie University. The Unit then extracted the administrative data using the encrypted numbers (PHRU data files contain only encrypted identifiers). In Ontario, the Institute for Clinical and Evaluative Sciences (ICES) conducted the linkage and the data extraction internally. ICES has broader access to the data compared to the other centers, but in a highly secure environment.

Derived Variables Measuring Health Services Utilization

Before proceeding with the analysis, derived variables were constructed from the administrative data. For example, for each NPHS respondent, annual usage rates for different types of health services were computed from that respondent's administrative data. Computer programs were developed and tested to compute these derived variables, and will be implemented in the regional offices.

Data Security and Access

Data for this project were accessed through the cooperation of provincial health departments and Statistics Canada, and as such were subject to provincial policies regarding data access and confidentiality, the Statistics Act and Statistics Canada's internal security policies. Researchers, and all staff accessing the pooled data, underwent security checks and were sworn in as deemed employees of Statistics Canada. As such, they were subject to the rules and regulations governing Statistics Canada employees, and could only access the linked data on Statistics Canada premises.

There are a number of carefully monitored security precautions that govern the access to data in Statistics Canada's regional offices. The data are housed in physically secure environments, and are not accessible through the Internet. Researchers are not allowed to remove data, or printouts containing data, from the regional offices unless they are at a sufficient level of aggregation for release. During the summer of 2001, changes in security requirements at regional offices necessitated additional delays, and in the short term resulted in the centralization of preliminary analysis at the Halifax regional office.

Lessons and Recommendations

What is the Value of this Data Linkage?

The linked data for this project combine the best information available on the demographic, socioeconomic, household, employment and health characteristics of Canadians (the NPHS) with the best and most detailed information available on health care utilization (administrative health care databases). As such, the data provide unique information on health determinants, health care utilization and health outcomes. In addition, these data are longitudinal, thus allowing researchers to examine processes that operate over time and to study determinants and outcomes. Also, because of the sampling design of the NPHS, the data assembled for this study were population-based, which allowed researchers to accurately describe trends and patterns in the population. With these benefits, the linked data enabled researchers to:

- Study the effects of health determinants on health services utilization and health outcomes.
- Study the effects of health events and health care utilization on factors such as employment, income and psychosocial characteristics.
- Make regional and inter-provincial comparisons.

Advantages and Disadvantages of the Project Organizational Structure

The organizational structure assembled for this project combined the best possible expertise for each of the data sources used. The project team was a collaboration between Statistics Canada, which has the primary expertise with the NPHS, provincial research teams who have extensive experience with the use of their provincial administrative databases, and the Canadian Institute for Health Information, which has a core competency in setting national standards for health data and resolving data compatibility issues. Although complex, such an organizational structure was necessary to bring the requisite skills together to undertake the research and to overcome barriers to integrating Canadian health information sources.

While the study team was well structured to address the research and technical aspects of this project, it was in a weak position to promote its broader goals to provincial governments and to federal agencies. The team did not initiate this study with “buy-in” from governments, especially provincial governments, to pursue its primary goal of assembling a federal/provincial organizational structure, and an integrated federal/provincial resource, that could be used for future research. In retrospect, such “buy-in” would have assisted the team in overcoming organizational barriers resulting from approval processes, contracts, and policies.

To contribute to the success of projects such as this in the future, federal/provincial/territorial priority setting mechanisms are needed to help coordinate support across levels of government. F/P/T committees may be best positioned to address this issue, and may be well positioned to recommend priority projects within their areas of emphasis.

The delays in this project caused by the diversity in provincial legislation, policies and procedures used to ensure the privacy of health information provide an excellent illustration of the need to harmonize federal and provincial privacy and confidentiality legislation, and associated regulations and review procedures. This issue is currently being addressed by the Federal/Provincial/Territorial Working Group on the Protection of Personal Health Information and is discussed in more detail below. There is also a need to facilitate the development of contractual arrangements and protocols for the sharing of health information for research purposes between the provinces, and between the provinces and Statistics Canada. These arrangements and protocols should include provisions supporting the role of academic research.

In summary, the most important organizational lesson learned from this project is that the promotion of successful federal/provincial health services and population health research efforts requires the assembly of federal/provincial/territorial research partnerships that:

- Draw expertise from the federal government, the provinces and academic research institutions; and that
- Link federal and provincial data resources.

These partnerships need to be augmented with federal/provincial/territorial organizational supports to:

- Set priorities to ensure support for similar projects at multiple levels of government.
- Harmonize review procedures, particularly around privacy and confidentiality.
- Develop policies and contractual arrangements governing the sharing of data between federal and provincial governments.

The Role of Statistics Canada

Even though provincial governments have access to the NPHS and in theory it would be possible to conduct a project such as this without the involvement of Statistics Canada, provincial contractual and security arrangements would make this hard to accomplish. Organizing data access through Statistics Canada offered a useful mechanism for sharing data between the federal and provincial governments, added credibility and security to the project and provided access to expertise. That arrangement did, however, introduce some substantial barriers and delays into the project as a result of the *On Premises Access Microdata Research Contract* that researchers had to sign establishing them as "deemed employees" of Statistics Canada. The version of the contract that was in place at the outset of this project required the "deemed employee" to produce a product for Statistics Canada, which implied that the primary study results had to be submitted to one of Statistics Canada's publications. While the contract included provisions on the submission of secondary analyses, or of primary results if the researchers were unable to meet the peer review requirements of Statistics Canada, to other publication venues, it also allowed Statistics Canada to restrict publication. Moreover, the editorial policies of Statistics

Canada publications impose limitations on the opinions and recommendations that researchers can express in reports, recommendations that may be important and acceptable in scientific journals but are inappropriate for a federal statistical agency.

The publication requirements in the original contract introduced two important barriers to this project. First, universities and research hospitals generally have firm policies requiring that research contracts guarantee publications rights. This is a key principle of academic freedom. Following the much-publicized case of Dr. Nancy Oliveri, a researcher at the Hospital for Sick Children who published results in violation of publication restrictions in a contract she had signed with a pharmaceutical company, this issue has become particularly important to universities and research hospitals. A number of the researchers on this project were unwilling to sign the contract because of this issue. To address the problem and avoid similar problems in future contracts, Statistics Canada made changes to the standard contract removing the publication restrictions, while still requiring initial submission to a Statistics Canada publication.

A second barrier was that submission of a report to the Health Transition Fund would be in violation of the contract. According to the contract, only secondary analyses could be included in a report to the Health Transition Fund, and those would have to be submitted following the publication of primary results in a Statistics Canada publication. For this study, it was agreed that empirical results would be simultaneously submitted to a Statistics Canada publication and to the Health Transition Fund, and that only results that did not pertain to the data (lessons learned, recommendations, etc.) could be disseminated by the Health Transition Fund. Restrictions on the submission of primary results, however, are likely to continue to be a barrier in the future, especially in projects where the primary project funding comes from a source other than Statistics Canada, and where the funding agency has its own reporting requirements. To avoid hampering its ability to participate in research partnerships, Statistics Canada should explore ways to make its publication requirements less restrictive by broadening its definition of the "product" required from deemed employees. For example, where Statistics Canada is an official partner in research initiatives, perhaps the "product" could be defined as Statistics Canada's deliverable to the partnership (e.g. a report released by another federal agency, such as Health Canada, with appropriate acknowledgement of Statistics Canada's contribution).

The Role of University-Based Research Centers and Institutes

Presently, most provinces have research centers and institutes with considerable expertise in the application of provincial health data to research. British Columbia, Manitoba, Ontario and Nova Scotia have established academically based research groups that act as custodians of provincial health data for research purposes and provide research expertise and technical expertise around data quality and organization. The emergence of research funding sources such as the Health Transition Fund, the Canadian Population Health Initiative, the Canadian Health Services Research Foundation, and the broadened mandate of the Canadian Institutes of Health Research, have helped to facilitate collaboration between academic research groups and between academic groups and federal/provincial/territorial governments.

The Role of the Canadian Institute for Health Information (CIHI) and the Canadian Population Health Initiative (CPHI)

The Role of CIHI

CIHI plays a critical role in the setting of data standards to promote compatibility between provincial health data sources. This project highlighted the need to combine provincial health data sources, but also demonstrated the complexity of doing so. The setting of standards for the coding of hospital discharge data greatly facilitated the use of hospital data for this project. Coding differences between physician claims databases, on the other hand, introduced substantial barriers.

The development of data standards as well as solutions to coding differences should continue to be high priorities for CIHI. In particular, there is a need to establish comparable data collection standards and protocols for different types of claims data and health information to avert the risk that a variety of different standards will be developed. CIHI has been approached by a number of groups, including the Association for Claims Exchange, the Canadian Pharmacists Association, and the Ministry of Health in British Columbia, with requests to assist in the coordination of governance structures and funding models to support the development of a national electronic claims standard. In response, it has launched the National e-Claims Standard Initiative (NeCST). NeCST's goals are to:

- Develop a common standards framework which encompasses all encounter/claims information.
- Develop and utilize a governance model that is simple and consensus-driven with balanced representation.
- Establish a funding base and funding model to ensure continued momentum for NeCST. At present, this initiative lacks a funding base.

CIHI and the federal government should also place a high priority on the broad dissemination of technologies that will facilitate inter-provincial and longitudinal health research. For example, the National Grouping System, developed by CIHI for the purpose of constructing the National Physician Database, could provide a valuable tool for inter-provincial studies. This resource, however, is not currently structured for dissemination to provincial research groups, so the dissemination of its methodology should be a high priority. In addition, CIHI and the federal government should help to ensure that provincial research centers have the capacity to implement Case-Mix Grouping. That the majority of research groups in this study did not have that capacity, largely because the costs were prohibitive, is a concern.

The Role of CPHI

This project was launched as a demonstration project for (CPHI) and was funded by the federal government as part of the Health Roadmap Initiative. The role of CPHI is to support research advancing an understanding of the determinants of the health of the Canadian population and to enhance the articulation of policy options to improve population health and reduce health inequities. Among CPHI's goals are to:

- Build research capacity in population health science and scholarship that complements investments by other funding agencies, and
- Contribute to the development of population health information systems.

This project helped to illustrate how these goals can be achieved. At the same time, it highlighted many barriers to the development of national infrastructure. CPHI should consider its role in brokering strategies to reduce these barriers.

Ethics, Privacy and Confidentiality Review: The Need for Harmonization

Clearly, the protection of privacy is a fundamental public concern that needs to be addressed through effective legislation and policy. Such legislation and policy must ensure strong protection for personal health information while providing access to information that is necessary for the evaluation of health policy, for the monitoring of the performance of the health care system, and for research with clear public benefits. In this project, it was not privacy legislation and policies per se, but the diversity in privacy legislation, policies and procedures, that imposed substantial barriers (For an overview of federal and provincial privacy legislation, see: http://www.cihr-irsc.gc.ca/ publications/ ethics/privacy/ compendium_e.pdf). It is noteworthy that no two reviews raised the same concerns, and that some reviews were much more restrictive than others. In some provinces, it was not clear what criteria were used for review, while in others (for example British Columbia and Manitoba) the criteria were clearly documented.

Existing legislation, and legislation under development, may help to introduce some common criteria for privacy and confidentiality reviews. The new federal privacy legislation, the Personal Information Protection and Electronic Documents Act (S.C. 2000, c. 5), applies to personal health information as of January 1, 2002, and will extend to activities operating entirely within provincial borders as of January 1, 2004. However, the act will not apply to provincial activities if provinces enact substantially similar legislation by 2004. As a result, most provinces are likely to enact new legislation in the next few years if they have not already done so. The new federal legislation applies to commercial activity, and thus would probably not apply to this study. Nevertheless, legislative activity and associated regulations are likely to impact the review procedures generally.

Regardless of the impact of new legislation, it is critical that provinces and the federal government work together to harmonize privacy legislation, as well as associated regulations and review procedures. Several ongoing activities may assist in this task. The Federal/Provincial/Territorial Working Group on the Protection of Personal Health Information has been working on reaching a Harmonization Resolution between all

Ministers of Health. Under this resolution, the Ministers would commit to a process of reviewing the protection of personal health information in their respective jurisdictions in an effort to develop a harmonized policy framework across the country. The Canadian Institutes of Health Research is engaged in a number of activities to help inform privacy legislation and ensure that new legislation is practical from a research perspective.

In sharp contrast to confidentiality reviews, ethics reviews imposed fewer barriers, and the review process was more systematic. In part, this was because a consistent set of standards for ethical reviews was used in all five provinces: the 1998 Tri-Council Policy Statement, "Ethical Conduct for Research Involving Human Subjects." The standardization was the result of a collaboration of the three major national research funding councils (The Medical Research Council (now the Canadian Institutes of Health Research), the Social Sciences and Humanities Research Council, and the Natural Sciences and Engineering Research Council). In order to receive research funding from either of these granting councils, research institutions must adhere to the Policy Statement. There would be considerable benefits to creating a parallel system for privacy and confidentiality reviews.

Priority Setting and Process: A Framework for Promoting Federal/Provincial Research Infrastructure and Initiatives

Overall, this project was viewed by gatekeepers as just one of many research projects seeking access to various sources of health data and was not given high priority. For the most part, the required review procedures were geared to the evaluation of the ethical issues and data requests pertinent to the specific research objectives of the project and were not set up to consider its broader infrastructure goals.

Where there were established review procedures and criteria, reviews generally proceeded smoothly regardless of gatekeeper priorities. However, in cases where special approvals were required, policy changes were needed, or the project required time inputs from third parties. The priority given to the project by a gatekeeper was an important factor, and low priority was often the cause of long delays, such as in the case of Ontario.

Given the growing need for timely information and research to assist health policy, health care planning and evaluation, the current environment where federal/provincial/territorial research efforts require years of special negotiation and approvals is problematic. For similar research efforts to proceed expeditiously in the future, a framework needs to be put in place to approve and support priority projects. Specifically, two improvements are needed. First, federal, provincial and territorial governments need a process designed to identify priority initiatives. It may well be that the Federal/Provincial/Territorial Advisory Committees reporting to the Conference of Deputy Ministers of Health are best positioned to assume this role. Second, clear procedures and policies need to be established to support priority projects. There needs to be mechanisms to link priorities to actual instrumental support, including provisions for the sharing of data and information between provinces and levels of government.

Literature Cited

Smith, R., and Haynes, S., (1992). Barriers to screening for breast cancer. *Cancer, 69*, 1968-1978.

Statistics Canada. National Population Health Survey general information and products and services. http://www.statcan.ca. Statistics Canada, 2001

Swain, L., Catlin, C., and Beaudet, M.P., (1999). The National Population Health Survey - its longitudinal nature. *Health Rep.*, *10,* 69-82.

Appendix

Required Approvals and Associated Barriers

Required Approvals and Associated Barriers

British Columbia

University Ethics: Yes

Privacy and Confidentiality

1. Data for the project were derived from the British Columbia Linked Health Database. The database contains a variety of health related data, is only used for research purposes, and resides at the University of British Columbia.
2. Approval to access the data was required from the Data Access Committee, composed of 'data stewards' for each of component database. The main purpose of the DAC is to ensure that research conducted is consistent with British Columbia's Freedom of Information and Protection of Privacy Act.
3. The data access committee required a specific, detailed specification of the data to be linked.[a]
4. The committee would not allow the release of potentially identifying information including 6-digit postal code.

Ministerial Approval

Ministry of Health officials are key members of the Data Access Committee, including members from the Ministry's Information and Privacy Branch and the Research and Evaluation Branch.

Other

The British Columbia team required clarification from Statistics Canada on policy for future use of the data.

[a] Initially, we decided to link all useful, non-identifying information in the administrative data. One goal of the project was to create an ongoing resource for research using the NPHS data, based on respondents consent for linkage. Moreover, because obtaining provincial reviews can be time consuming, we wanted to avoid decisions on data requirements at the outset that might be limiting later on. However, many review bodies are geared to providing data on a project-by-project basis, and thus require detailed specifics on the data requirements for the project, along with the rational for requested data elements. In some cases, grouping of variables (e.g. age) was required.

Saskatchewan

University Ethics: Yes

Privacy and Confidentiality

1. Reviewed by the Data Access Review Committee internal to the Department of Health. The Committee, which has delegated authority, reviews all data request for compliance with the Department's data access and confidentiality policies. The Committee does not include academic representation.
2. Saskatchewan Health's data release policies and practices can be quite restrictive about data sent outside the Department.
3. Saskatchewan Health required very detailed specifications on the data to be sent, and imposed a number of restrictions. For example, they would not provide detailed ICD codes required for CMG groupings from the hospital data. Extensive grouping of fields was required.
4. Saskatchewan Health only agreed to allow linkage and release of data for those persons who had given consent in both waves of the NPHS (i.e. the longitudinal file). Accordingly, we were unable to use data for NPHS respondents who were lost to follow-up between waves of the survey (in this case less than 50 subjects).

Ministerial Approval

The Data Access Review Committee internal to the Department of Health has delegated authority.

Other

Cost to access data was initially prohibitive. However, in the end they extracted a limited set of administrative data at low cost.

Data was released to the provincial investigator who then signed a contract with Saskatchewan Health.

Manitoba

University Ethics: Yes

Privacy and Confidentiality

1. Reviewed by an independent privacy and confidentiality committee.
2. The committee required a specific, detailed list of the fields to be sent to Statistics Canada.[a.]
3. Required that confidentiality contracts be signed by all project investigators.
4. The final approval was inflexible to change in protocol. For example, the original approval would not permit hospital data to go to CIHI for coding/grouping.

Ministerial Approval

Essentially—linkage could not be conducted at MCHPE, so Ministry support was required.

Other

1. Required clarification from Statistics Canada on future use of the data.
2. Linkage had to be conducted at Manitoba Health rather than MCHPE. Thus, the speed of getting the project done was subject to Manitoba Health's priorities rather than that of the research team. This has created problems for refinements to the 1994 link.

[a.] Initially, we decided to link all useful, non-identifying information in the administrative data. One goal of the project was to create an ongoing resource for research using the NPHS data, based on respondents consent for linkage. Moreover, because obtaining provincial reviews can be time consuming, we wanted to avoid decisions on data requirements at the outset that might be limiting later on. However, many review bodies are geared to providing data on a project-by-project basis, and thus require detailed specifics on the data requirements for the project, along with the rational for requested data elements. In some cases, grouping of variables (e.g. age) was required.

Ontario

University Ethics: Yes

Privacy and Confidentiality

Review was conducted within ICES by a privacy officer and the CEO. In addition the legal branch reviewed the project for compliance with the provincial Freedom of Information and Protection of Privacy Act. ICES has considerable latitude on access to and use of the NPHS share data internally. For example, at the point of project initiation, they were already active in linking the NPHS to administrative data.

Ministerial Approval

1. ICES is not authorized to send data to other organizations without ministerial approval.
2. Obtaining ministerial approval was very difficult largely because the project is not a priority. Approval was granted in Summer 2001.
3. Now that ministerial approval has been obtained, approval from the legal branch is required. This is still pending.

Other

The privacy commissioner requires that the data must have a specified destruction date.

Nova Scotia

University Ethics: Yes

Privacy and Confidentiality

Review was conducted within the Population Health Research Unit (PHRU) according to pre-approved criteria.

Ministerial Approval

PHRUs contract with the Nova Scotia Dept of Health does not authorize it to send data out of the province. Thus, ministerial approval was required and obtained from the Deputy Minister.

Other

PHRU works only with encrypted identifiers, and is dependent on Atlantic Blue Cross Care to implement encryption for linkage.

Statistics Canada

University Ethics: No

Privacy and Confidentiality

Approval from the Statistics Canada Policy Committee was required.

Ministerial Approval

No

Other

Access to data is subject to the Statistics Act, as well as to the provisions of the Deemed Employee contract. The deemed employee contract included problematic requirements on (a) publication and (b) submission of reports to other funders. This project required substantial changes to the deemed employee contract.