



Education

Éducation

Dr. Kazanjian is Associate Director of the Centre for Health Services and Policy Research, and Associate Professor in the Department of Health Care and Epidemiology, University of British Columbia, and a Faculty Associate with the BC Centre of Excellence for Women's Health, BC Children and Women's Hospital and Health Centre, Vancouver, BC.

This article has been peer reviewed.

CMAJ 1998;159:342-5

Understanding women's health through data development and data linkage: implications for research and policy

Arminée Kazanjian, DrSoc

Women's health has traditionally been conceptualized from a specialty-focused definition, grounded in the biomedical model. Fragmentation has also existed at the level of health care delivery: medical and mental health services have usually been provided by generalists and reproductive services by specialists. Yet research on women's health has broadened appreciably from its early emphasis on reproduction. A wide variety of issues are now being raised and studied, from the effects of the menstrual cycle on diseases and treatments to the health of older women.

More sophisticated and complex conceptualizations of women's health have now evolved, grounded in a biopsychosocial model: women are more than an aggregation of cells, body parts, social actions and social statuses.¹ There is some evidence that the difference in morbidity rates between women and men may be largely explained by nonbiologic factors, including socioeconomic and psychosocial variables. Developing more inclusive models of health requires recognizing and examining the complexities and diversities in women's lives, a vast array of social and cultural identities that define the roles, status and experiences shaping women's health.

Feminist perspectives have spurred recognition of how gender affects the causes, natural history and treatment of disease. Feminist research posits that women's health problems must be understood as socially, culturally and economically produced, not as isolated, individual, biological events that can be explained outside the contexts in which they emerge.¹ This perspective highlights the underlying social dynamics of what actually produces health for different groups of women.

In conjunction with the variety of perspectives on women's health, several frameworks have been developed to reflect different approaches to understanding and studying this topic. Such frameworks provide overviews of the information required to undertake research in women's health. The purpose of this article is to describe, in some detail, the level and type of information available through existing databases that may be further developed by linking 2 or more such databases. Thus, the discussion of selected models and frameworks serves only to illustrate issues pertaining to database development and analysis. The intent is to demonstrate that the secondary analysis of existing databases, using any of the models or frameworks currently under consideration, is a valuable tool for the study of women's health. In addition, conceptual frameworks can be revised, refined or rejected through empirical analyses using existing databases.

Women's health research and policy frameworks

Ten years ago the Conference of Deputy Ministers of Health convened a Federal/Provincial/Territorial Working Group on Women's Health. A report was produced describing a framework for the development of policies and programs.² A list of issues pertaining to the health of Canadian women, based on the international literature on this subject, was identified: mental health, reproductive health, cancers occurring in women, occupational and environmental health, nutrition and active living. The report emphasized the diversity of female populations, a phenomenon that, ironically, seems to relegate large numbers of women to minority status.

At about the same time, the US Public Health Service defined women's health to include screening, diagnosis and management of conditions that are unique to



women, are more common in women, are more serious in women and have manifestations, risk factors or interventions that are different for women.³

More recently, a framework was proposed by Chesney and Ozer⁴ to include 7 content areas: reproductive health, diseases more common in women than men, leading causes of death among women, gender influences on health risk, societal influences on women's health, violence against women, and women and health care policy.

Provincial and national databases for the study of women's health

Appreciable public resources have been invested at the provincial and federal levels to develop and maintain large computerized databases. These are of 2 major types, depending on their purpose: administrative databases and survey databases (Table 1). Each has some strengths and limitations with regard to gender analysis and research on women's health. However, these information resources remain largely underutilized and often unknown to researchers, planners and policy-makers. Although information technology — hardware and software — continues to develop at an impressive rate, its true value cannot be fully realized in health research until better use is made of existing data, vast sources of knowledge that can be tapped at reasonable cost with no response burden for the population.

Administrative databases

Administrative databases include computerized systems supporting provincial health care and social services programs (Table 1). The central administration of these publicly funded programs permits each province to develop a set of

population-based administrative health service data files. Although they were created to support provincial payment systems and management functions, they contain a large array of data for entire populations and provide excellent opportunities for important research on women's health. The data can be used to guide and evaluate existing programs and policies and to inform and plan such activities in the future.

Data on physician payments, hospital admissions and discharges, long-term care and continuing care, pharmaceutical and home care can be used to enhance our understanding of women's health needs and of how patterns of health care utilization are tied to various aspects of women's lives (for example, age and geography). Although these administrative data are collected at the provincial level, they are developed and maintained as national databases through the Canadian Institute for Health Information, which facilitates interprovincial and national analyses.

In addition to program delivery and management support databases, archival data that document births, deaths, stillbirths, marriages and divorces (vital statistics) along with disease registries account for other administrative databases. Disease registries are usually maintained at the provincial level, although the Canadian Cancer Registry has existed since 1992 (and in a different format since 1969). These data provide valuable opportunities for research that focuses on measuring health status and outcomes and examining differences between the sexes in morbidity and mortality.

Finally, census data, available from 1971, provide comparable electronic data on individuals, families and households every 5 years (every 10 years since 1871; in 1956 the shorter census was introduced). During each census basic demographic information is collected for all persons in a defined geographic area. More detailed social, cultural and economic information is collected from a 20% sample of the population.

Survey data

The second group of databases is based on surveys (Table 1). A multitude of survey data are available: longitudinal cohort surveys, time-series surveys and ad hoc surveys. Longitudinal cohort surveys (sometimes referred to as panel surveys) follow a panel of respondents for 2 or more times (with waves of questionnaires or interviews). A common core of questions may be supplemented with additional ones for subsequent waves.

Another approach is the time-series survey, sometimes referred to as a cross-sectional survey. It usually involves the administration of a core set of questions to representative samples of individuals from the population at different points in time. As in the longitudinal survey, core and supplemental questions are posed; however, different individuals constitute the sample each time.

Table 1: Types of databases

Administrative data		
Programs	Archives	
Provincial payment systems, e.g., fee-for-service payments to physicians	Disease registries, e.g., cancer, congenital diseases	
Payments to clients in long-term care	Census statistics, e.g., demographic, social, economic	
Management functions, e.g., births, deaths, marriages		
Survey data		
Longitudinal	Time-series	Ad hoc
Same panel of respondents over time, e.g., National Population Health Survey	Representative samples at different points in time, e.g., General Social Survey	One-time samples on special topics, e.g., Aboriginal Peoples Survey



Ad hoc surveys are one-time efforts, usually in response to a specific information need.

In Canada the General Social Survey (GSS) is an annual cross-sectional survey; key topics repeated every 5 years have included health (1985 and 1991) and social support (1985, 1990 and 1996), among others. The National Population Health Survey (NPHS), a panel survey, was first undertaken in 1994 and was repeated in 1996; detailed information was sought on sociodemographic variables, health status, stress, social resources, mental health and health promotion. These surveys provide a wealth of information on women, their lives and health. An ad hoc (or special) survey on Violence Against Women was undertaken in 1993; the Canadian Health and Disability Survey in 1983/84 was another special survey with particular relevance to women's health.

These survey databases contain detailed information on demographic, socioeconomic and psychosocial variables, in addition to information about health and wellness (Table 2). Detailed descriptions of Canadian surveys and the variables they cover are available elsewhere.⁵ In contrast to the administrative databases, which may be more pertinent to the analysis of women's access to care, their use of services and the system's response to women's health care needs, survey databases permit the analysis of the social determinants of health and explore the gender-related aspects of such determinants. However, unless the surveys are designed at the outset to incorporate questions relevant to women's lives, their well-being and health, secondary analyses will not yield useful information.

Because they contain self-reported data, surveys have some limitations in terms of the validity and reliability of information, whereas administrative data limitations pertain to the narrow range of information available. Both

types of limitations are inherent attributes of data collection designs that should be acknowledged in the analyses; however, they should not preclude the use of the data.

Federal, provincial and territorial systems of collecting and developing population-based data enable multiple data sources to be used through data linkage.⁶ The capacity to link population-based health data was first developed in Manitoba⁷ and more recently in British Columbia.⁸ Linked databases allow the tracking of an individual's episodes of care across programs (e.g., physician's office visits, hospital in-patient care, surgical day-care, home nursing care, long-term care), to provide a more comprehensive analysis of population health needs, and across time (for example 1986–1996), to enable longitudinal analyses of trends and patterns of care.

Future linkage

It is possible to link the administrative databases retrospectively with survey data to enhance the analytic value of the variables in each type of database and to allow the application of broader frameworks for understanding and examining women's health. For example, the sixth cycle of the GSS was the health survey conducted in 1991, a census year. Linking individual GSS records, which contain information on health and social support, with those of the 1991 census and further with those of provincial medical and hospital utilization databases for multiple years would provide valuable new information at a reasonable cost. Researchers could then study the effect of social support on men and women, controlling for various sociodemographic factors available in the census data and measuring its influence in health care utilization on the basis of the administrative databases. However, this type of study is not currently done,

Table 2: Characteristics of various types of data and type of analysis relevant to women's health research

Data characteristics	Type of analysis
Administrative data	
Population-based	Compare patterns of service utilization by key variables, e.g., rates of cesarean section by health region
Individual as unit of observation	
Historic	
Retrospective	Identify episodes of care by subpopulations, e.g., treatment of depression among adolescent girls
Survey data	
Fixed cohorts and waves or representative samples	Study life-cycle transitions, e.g., family structure and perceived health
Individual or household as unit of observation	
Multiple waves	Introduce new subjects reflecting new concerns, e.g., physical fitness during perimenopausal years and future fractures
Core topic with or without supplemental topics	
Prospective	
Linked data	
Scrambled unique identifiers	Evaluate social and health policy interventions, e.g., effect of income subsidy on health care utilization and health outcomes and perceived health
Prospective and retrospective	
Combined subjective and objective information	Explore new phenomena, e.g., aspects of determinants of health according to sex



partly because of possible barriers to data access. Data repositories and their stewardship by various government agencies are still under discussion and negotiation. Regrettably, these debates usually occur without the benefit of input from researchers in the field of women's health.

Prospective linkages of archival and survey data are less difficult and more accurate through the use of unique identifiers such as the personal identification or health insurance number issued to all individuals covered under our universal health insurance schemes. The NPHS data contain that information for the overwhelming proportion (90%) of respondents who consented to the use of their identification numbers for data-linkage purposes. This consent for the use of such identifiers contrasts with concerns raised by many individuals regarding data confidentiality by agencies that undertake surveys or maintain archival data. Increased vigilance over time about data confidentiality and security are usual for research organizations engaged in data linkage. Standard procedures that are subject to audits include security arrangements about where data will be held and processed, assurances that no data on individuals will be released in public documents, including the signing of an "oath of confidentiality," and an indication that the data will be returned or destroyed upon project completion. Such safeguards ensure that data on individuals are handled by the minimum number of systems operators and that unique identifiers are scrambled or that individual data are aggregated to ensure anonymity.

Data linkage provides a powerful analytic tool for describing and measuring the gender-related aspects of key social determinants constructed through the analysis of numerous surveys on health and illness behaviour, attitudes, values and beliefs, as well as sociodemographic information on individuals, families and communities.

Clinical studies and future prospects

It is now well known that, historically, clinical trials have excluded women as subjects, yet the clinical care of women is based on the findings from such trials. In heart disease, for example, the MRFIT⁹ studies examined mortality in 12 866 men only; the Physicians' Health Study¹⁰ looked at low-dose aspirin therapy and the risk of myocardial infarction in 22 071 men; the Health Professions Follow-up Study¹¹ explored the association between coffee consumption and heart disease in 45 589 men. The exclusion of women from studies is not confined to clinical research nor to the United States. The Ontario Longitudinal Study of Aging has followed annually, since 1959, a cohort of elderly men only. However, the US Nurses' Health Study (and subsequent analyses) has followed a cohort of 95 000 nurses since 1976 to partly offset the imbalance.¹²

A recent major effort in the US to describe the ethical

and legal issues related to the inclusion of women in clinical studies culminated in a 2-volume report.¹³ The paucity of information on women's participation in clinical research in general prompted a recommendation by the authors of that report to support the National Institutes of Health Revitalization Act of 1993 and a reiteration of the importance of establishing a central registry of clinical studies that includes information on sex, race and ethnicity as well as comprehensive information on results. Although there are a number of clinical trial registries throughout the world, most focus on specific medical conditions, and information about the sex of participants may or may not be available.

A central registry of all Canadian clinical studies (including international studies with Canadian participation) incorporating information about the sex of participants is possible through the Canadian Cochrane Network and Centre, part of the international Cochrane Collaboration.¹⁴ This registry could be part of a sophisticated information system concept — currently under discussion by federal government officials — consolidating major health databases such as one on the First Nations, a national health surveillance system and so forth, all containing data relevant to the study of women's health.

References

1. Ruzek SB, Clarke AE, Olesen VL. Social, biomedical and feminist models of women's health. In: Ruzek SB, Olesen VL, Clarke AE, editors. *Women's health*. Columbus: Ohio State University Press; 1997. p. 11–28.
2. Canada Federal/Provincial/Territorial Working Group on Women's Health. *Working together for women's health: a framework for the development of policies and programs*. Ottawa: Health Canada; 1990.
3. Office on Women's Health. *PHS action plan for women's health*. Pub no (PHS) 91-50214. Washington: US Public Health Service; 1991.
4. Chesney MA, Ozer EM. Women and health in search of a paradigm. *Women's Health Res Gender Behav Pol* 1995;1:3–26
5. *Finding data on women: a guide to the major sources at Statistics Canada*. Ottawa: Status of Women Canada; 1998.
6. Newcombe HB. *Handbook of record linkage*. New York: Oxford University Press; 1988.
7. Roos NP, Shapiro E, editors. Experience with a population-based health information system. *Med Care* 1995;33(12 Suppl).
8. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. *Can J Public Health* 1998;89(4):270–3.
9. Multiple Risk-Factor Intervention Trial Group. Statistical design considerations in the NHLI multiple risk factor intervention trial (MRFIT). *J Chron Dis* 1977;30:261–75.
10. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med* 1989;321(3):129–35.
11. Grobbee DE, Rimm EB, Giovannucci E, Colditz G, Stampfer M, Willett W. Coffee, caffeine, and cardiovascular disease in men. *N Engl J Med* 1990; 323(15):1026–32.
12. Michels KB, Rosner BA, Manson JE, Stampfer MJ, Walker AM, Willnett WC, et al. Prospective study of calcium channel blocker use, cardiovascular disease, and total mortality among hypertensive women: the Nurses' Health Study. *Circulation* 1998;97(16):1540–8.
13. Mastroianna AC, Faden R, Federman D, editors. *Women and health research*. 2 vol. Washington: National Academy Press; 1994.
14. Haynes B, Jadad A. *Canadian Cochrane Network and Centre 1997 annual report*. Hamilton (ON): Health Information Research Unit, McMaster University; 1998.

Reprint requests to: Dr. Arminée Kazanjian, Associate Director, Centre for Health Services and Policy Research, University of British Columbia, S184 - Koerner Pavilion, 2211 Wesbrook Mall, Vancouver BC V6T 2B5; fax 604 822-5690; arminee@chspr.ubc.ca