A DISCUSSION PAPER
ON OUTCOMES AND MEASUREMENT
IN THE VOLUNTARY HEALTH SECTOR IN CANADA

Prepared for the

Voluntary Health Sector Project

October 1999

Barbara Legowski
Terry Albert

Foreword

Voluntary organizations are receiving increasing recognition for the role that they play in Canadian life. Although their contributions are evident in most communities, they have yet to be systematically documented. In 1999, the Canadian Centre for Philanthropy, Canadian Policy Research Networks, Health Canada and the Coalition of National Voluntary Organizations undertook a joint initiative to enable researchers to begin documenting the contributions of voluntary organizations working in the area of health. Two papers were commissioned to lay the foundation for further empirical studies in this area. *Developing a Typology of the Voluntary Health Sector in Canada: Definition and Classification Issues* was prepared to address the important issue of defining what organizations should be included in such studies and to develop an appropriate system for classifying these organizations. *A Discussion Paper on Outcomes and Measurement in the Voluntary Health Sector in Canada* was prepared to provide guidance about how to measure the economic and social contributions of voluntary health organizations. It is hoped that these two companion papers will help set the stage for further research into the social and economic value of voluntary health organizations in Canada. Copies of both papers and of their respective Executive Summaries can be obtained by contacting any of the respective partner organizations.

The Management Committee for this project would like to take this opportunity to thank the many people who made contributions to these studies. Individuals from a number of key organizations contributed their time and insights through discussions with the authors of the papers and we thank them for their valuable input. We also thank the five anonymous peer reviewers for their thoughtful comments and suggestions. In the final analysis, however, the papers reflect the views and opinions of the authors and any errors of omission or interpretation are their own.

Special thanks are due to Health Canada for financial support of the project.

Members of the Voluntary Health Sector Project Management Committee

Catherine Adam, Health Canada
Terry Albert, Canadian Policy Research Networks
Karl Benne, Health Canada
Katie Davidman, Canadian Policy Research Networks
Angela Febbraro, Canadian Centre for Philanthropy
Michael Hall, Canadian Centre for Philanthropy
Rhonda Hynds, Coalition of National Voluntary Organizations
Mary Jane Lipkin, Health Canada
Graham Lowe, Canadian Policy Research Networks
Penny Marrett, Coalition of National Voluntary Organizations
Kathryn McMullen, Canadian Policy Research Networks
Jim O'Brien, Canadian Diabetes Association

October 1999

TABLE OF CONTENTS

SYNOPSIS

Introduction and Background

The Canadian Policy Research Networks, the Canadian Centre for Philanthropy (CCP), Health Canada and the Coalition of National Voluntary Organizations have entered into a partnership to conduct a study on the voluntary health sector. The object of the project, entitled The Voluntary Health Sector: An Assessment of its Size, Scope and Value, is to assess the contribution of the voluntary health sector by:

- giving the sector form and definition through boundary mapping and classification;
- analyzing and measuring both the qualitative and quantitative economic and non-economic contributions of the sector; and
- ensuring the dissemination of research findings to a variety of audiences, including the voluntary health sector, government, researchers and other interested parties.

The *Discussion Paper on Outcomes and Measurement in the Voluntary Health Sector*, which satisfies the project's second objective, is summarized below. The paper's objectives are:

- to review the literature on measuring health outcomes;
- to consider the feasibility of measuring the inputs, outputs and outcomes of health nonprofits; and
- to make recommendations on what can be measured to determine the contributions of health nonprofits to Canadians' health and/or to the health care system.

The data needed to derive the value of the voluntary health sector may be collected through survey and/or case studies.

Defining the Voluntary Health Sector

The research for this paper was undertaken without predefined boundaries or a classification for a voluntary health sector (VHS). Therefore, a working definition was developed for the sector to give it scope for the duration of the research. The process of definition had two steps: the first identified the health organizations or organizations working in health in such a sector; the second identified which of these organizations' activities had a health outcome.

Two existing organization-based classification systems were first considered: that of Revenue Canada and that of the International Classification of Nonprofit Organizations (ICNPO). Holistic definitions of health and the population health paradigm were also considered in delineating the sector. These had the potential to significantly broaden the scope of a voluntary health sector.

The Revenue Canada classification of registered charities and unregistered nonprofits whose stated purpose is health was selected over the ICNPO health category even though at first glance there appears to be some overlap and ambiguity in the Revenue Canada categories[1]. The ICNPO appeared too institutionally selective while Revenue Canada's classification is broader. The very

broad scope suggested by a population health paradigm was rejected 1) because of practical difficulties experienced by the field in relating program outcomes to population health determinants, and 2) because data collected to date describing potentially a voluntary health sector by, for example, Revenue Canada, had been limited to those organizations that recognized themselves as involved in health. As well, it seemed presumptuous to allocate organizations to a health sector that had not themselves identified their actions as directly contributing to health outcomes.

Next, an activity-based dimension with five domains was added to the Revenue Canada organization-based classification. The domains provide a means of relating discrete organizational processes to specific measurable health outcomes, where a health outcome is:

> a change in the health of an individual, a group of people or population that is attributable to an intervention or series of interventions (Rissel, Ward and Sainsbury, 1996).

Our working definition for a voluntary health sector was therefore:

> the group of formal organizations that are registered charities or nonprofit organizations without charitable status with a stated purpose or direct interest in health demonstrated through the following domains of activity that contribute to health outcomes – service provision as either disease and illness prevention/health promotion, and care; research; advocacy; fundraising; and practitioner regulation.

Conceptual Framework for Outcomes and Measurement

Having identified a voluntary health sector for our research, we combined the organizational activity domains defining the sector with a traditional inputs-through-outcomes model of organizational performance, where outcomes are the benefits for participants during and after program activities (Plantz, Greenway and Hendricks, 1997 - see Figure 1, p.8). Osborne and Tricker (1996) refer to this same model as the different parts of a service production process that make up a conceptual framework for studying organizational performance.

Not only do inputs through outcomes create the framework through which organizational performance can be measured, they also represent the spectrum across which value can be derived. There is inherent value that can be measured at each point in the spectrum (see Figure 2, p.9).

Combining the organizational activity domains with the spectrum of organizational value formed the basis for a conceptual framework for our research. We added to this a societal domain recognizing that products and services of nonprofits are returned to the community. They add value to social capital and cohesion, that is, add value that is more socially than privately distributed (The Steering Committee for the Social Accounting Framework Project, 1998 – subsequently referred to as The Steering Committee; Jensen, 1998; Veenstra, 1995). In this way, the conceptual framework for the research respects the standard model of how an organization can measure performance at the same time that it advances the discussion of the kinds of value or impact that can be attributed to the activities undertaken (See Table 1, p12).

To understand what the field is experiencing in terms of general measurement and outcomes measurement in particular, the paper explored briefly the performance measurement methods that evaluation science makes available to nonprofits. Historically, nonprofits have measured organizational and program-related inputs and outputs, largely driven by the reporting requirements of their funders (Plantz, Greenway and Hendricks, 1997). Examples of inputs are revenues, assets, paid staff, volunteers and volunteer time; examples of outputs are numbers of clients served and programs provided. More sophisticated measures of outputs add the dimension of quality of services provided, such as infection rates in hospitals. Generally, organizations have shown how much effort has been generated for how many people.

More recently, funders and the public are placing greater emphasis on value-for-money decisions for resource allocation which require program outcome information – they want to know if anyone is better off as a result of the services provided and they want to know the cost-benefit (The Panel on Accountability and Governance in the Voluntary Sector, 1999).

The shift away from traditional inputs and outputs has intuitive logic. In response, evaluation science is delivering an increasing array of methodologies for measuring not only program outcomes but also various other aspects of the program delivery process. Methods range from rigorously scientific experiments with control groups, to evaluations that deliberately involve many participants and collect numerous viewpoints in order to attach as large a constituency as possible to a particular program rendering a collateral benefit to the program operation.

Many issues arise with such a growing emphasis on outcomes measurement. For one, defining and measuring outcomes can be complex because outcomes have multiple dimensions. For example, the combination of possible measurement methods and time frames for measurement can create a number of different outcomes for any activity being evaluated. This is illustrated in Diagram 1(see p.21), where outcomes have three dimensions, placed along three axes. Along the *X* axis are the organizational and societal activity domains. Along the *Y* axis are direct and indirect outcomes, direct being those that are the ultimate intention of an activity and indirect being proxies for the ultimate outcomes or bridges to them. The *Z* axis represents outcomes determined by time. The time periods range from short to intermediate to long term.

Most outcome measurement focuses on the shorter terms. Evaluations of interventions that have a long lag time between the delivery of the service and the manifestation of an outcome are not common due to the complexities associated with longitudinal data collection, multi-year expenses and a largely funder driven bias for demonstrating short term or quick results.

Nonprofits' Self-evaluation and Measurement

The choice of evaluation methods is a continuing debate for nonprofits. They often find themselves providing complex social and health services, which in some respects is their niche (Hirshhorn, 1997; Webber, 1994). Outcomes for these services are inherently difficult to assess, or are too diffuse or broadly based to be readily measured with quantitative techniques that are favoured within a positivist frame of thinking. As well, nonprofits typically have a wide and

complex range of stakeholders, making the selection of what to measure, and how, a difficult decision especially given limited resources (Voluntary Sector Roundtable, 1997).

Viewpoint or perspective concerning the choice of measure is an important consideration. For example, some argue that what is perceived to be of value regarding an organization's services or products is subjective and negotiable, created by individuals involved in a specific context and capable of evolving as the individuals interact (Forbes, 1998). The evaluation literature encourages organizations to identify for themselves the outcomes or other aspects of program delivery for which evaluation results will optimize program effectiveness, as opposed to measuring only what is of interest to powerful stakeholders such as funders (Plantz, Greenway and Hendricks, 1997). Funders are in fact cautioned against imposing outcome measurement onto the voluntary sector without a full understanding of its limitations and possible negative impacts. For example, inappropriate linkage of funding to specific outcomes measurement can lead to: resources shifting from services to measurement with no direct benefit to programs; prevention and development programs with harder-to-measure outcomes being penalized; "creaming" being promoted (the deliberate selection of clients or programs most likely to have positive outcomes leading to increases in inequalities in health status); the discouraging of risk taking and innovation; and the fuelling of interagency competition as opposed to cooperation (Plantz, Greenway and Hendricks, 1997).

Along with the above hazards of an overemphasis on outcomes measurement are ethical issues surrounding information disclosure by nonprofits. Outcomes measurement delivers new information to the public domain creating the potential for linking, for example, inputs to outcomes (cost-benefit). Having in the public domain new information about a nonprofit's performance can jeopardize the value that an organization places on the privacy of certain clients or staff; the exposure may encourage organizations to adjust results to be more favourable; information may be misinterpreted if released or used out of context; and organizations may be placed at a disadvantage among their competitors, particularly for-profit agencies, that are not obliged to expose their performance results to the same degree.

While evaluation science is furnishing numerous approaches to performance measurement, the barriers to program evaluation by nonprofits are for the most part related to capacity, referring to lack of funding, the need for more training and assistance and the need for research-based comparative indicators (Zacharakis-Jutz and Gajenayake, 1994; Rodriguez-Spagnolo, 1992; Taylor and Sumariwalla, 1993). In some cases, organizations are reluctant to measure performance for fear of negative repercussions by funders if results are poor. Reluctance also stems on the one hand from a lack of direction from funders as to what evaluation results will be recognized, and, on the other hand, from a sense that funding decisions are more political than rational rendering even positive evaluations as no guarantee of funding.

Despite the reluctance, parent organizations and some funding agencies are encouraging and assisting smaller organizations to measure outcomes by offering training, workshops, pilot projects, and by providing measurement tools. Larger organizations have the advantage of greater resources so are likely ahead regarding measuring outcomes. Some, in fact, use positive outcomes as promotional tools to attract donations and other funds. Nonetheless, on the whole, in the absence of standardized measures of impact and with the perceived strong need for

nonprofit sector accountability, outcomes measurement is at this time reactive and ad hoc (The Steering Committee, 1998). Individual organizationally-defined outcomes can derive value, but lack a common unit of measure across organizations.

Hence, to address this dilemma, social accounting or social audit would be a measurement approach worthy of further examination. Using evaluation criteria and a participatory evaluation approach appropriate to nonprofit culture, the method attributes a comparative economic value to nonprofit activities and highlights their contributions in the social domain. Specifically, it delivers a social return-on-investment ratio per organization or program whose components may be additive across a voluntary health sector (The Steering Committee, 1998; Richmond, 1999). The evaluation criteria should be selected within the context of the emergent or social constructionist approach to effectiveness evaluation that has appeared in this literature over the past two to three years (Forbes, 1998). Again, more investigation is warranted.

Other common unit or universal measures of outcome in the health arena include measures such as the quality adjusted life years (QALYs) produced by specific interventions and programs. QALYs are generic composite measures that combine gains in length of life with gains in quality of life as opposed to disease-specific natural measures such as blood pressure, blood sugar, cholesterol, etc. As such, QALYs could theoretically be used to measure health outcomes produced by the programs of individual organizations and across organizations in a community or a province. The tool box of the health economist has four major approaches to economic appraisal, two of which can include QALYs in the denominator of the cost-effectiveness ratio where the costs of programs or interventions are related to consequences or outcomes (Drummond et al., 1987). The feasibility of employing this measure of health outcome should be considered within the research design phase of future evaluative research concerning the VHS.

Conclusions

The relevance of the paper to the major objectives of the larger study on the voluntary health sector boils down to the feasibility of the subsequent research design and methodology in determining the value of the sector. Presumably the unit of observation for the sector is the organization. The paper demonstrates that economic and otherwise quantitative value can be extracted discretely from organizational inputs and outputs. While the importance and relevance of measuring contributions to the organizational activity domains and the societal domain are undisputed, the first order priority may best be confined to some crude additive measures that can represent the whole sector. These could be inputs and outputs by activity domain.

The research design and methodological issues that remain unresolved for the larger voluntary health sector project have a major bearing on the selection of variables to be measured and, moreover, on how any data are brought to bear on the research objectives. For example, if the organization is the unit of observation, then it follows that the universe must next be defined – presumably the geographic universe is Canada. While we developed a working definition and scope for the voluntary health sector for this paper, the final typology will inform the actual boundaries of the sector. In the meantime, a key methodological issue is whether a statistical sample of voluntary health sector organizations will be needed. If so, different approaches need

discussion. For example, if a multi-pronged method were selected that focussed on the national and provincial level organizations and on a sample of cities (small, medium and large), then it might be possible to construct a composite picture and collect the measures of this geographical universe through extrapolation and other estimating techniques. Another approach may be a component modelling one, with the identification of "natural laboratories" such as Peterborough or Sherbrooke where the cities' compactness and generalizability have allowed corporate marketing studies and surveys. On the other hand, case studies would provide the opportunity for gaining a depth of understanding on organizational and programmatic inputs, outputs and outcomes.

In terms of what should be measured, the possible data elements for collection are in Table 2 (p. 40), which is a simplification of the original conceptual framework. A preliminary compilation and a sectoral consultation and peer review would be necessary prior to any launching of data collection. As well, the ethical issues described in the paper would require discussion.

Because the feasibility of collecting individual program outcomes data and contributions to a societal domain is low compared to input and output data, asking for a simpler version of these contributions is worth consideration. For example, volunteers in an organization could be asked for their personal perceptions of what volunteering has added to their lives according to their contributions by activity domain (for example, through fundraising). Finally, in terms of deriving a macro measure of social value, it is recommend that the social accounting method be investigated further and brought for consideration to sectoral consultations and peer review.

1 INTRODUCTION

The Canadian Policy Research Networks, the Canadian Centre for Philanthropy (CCP), Health Canada and the Coalition of National Voluntary Organizations are partners in a study of the voluntary health sector. The object of the study, entitled The Voluntary Health Sector: An Assessment of its Size, Scope and Value, is to determine the contribution of the voluntary health sector by:

- giving the sector form and definition through boundary mapping and classification;
- analyzing and measuring both the qualitative and quantitative economic and non-economic contributions of the sector; and
- ensuring the dissemination of research findings to a variety of audiences, including the voluntary health sector, government, researchers and other interested parties.

This paper is intended to partially address the second objective of the project through a literature review, informal interviews and exploratory thinking. It was prepared in isolation of a typology of the VHS. As such, an intentionally generic approach to outcome and activity measurement was adopted. In other words, instead of focusing exclusively on outcomes and other measures in the health arena, the paper allows for a blurred boundary between health and social services. Hence, the paper is a work in progress that should be updated once a typology for the VHS is developed. It is envisaged that this would also be accomplished through a substantive consultation process with the VHS. Ultimately, this initial work will inform the research design and methodological development for the empirical work on assessing the various dimensions of the value of the voluntary health sector in Canada.

2 OBJECTIVES OF THE PAPER

The specific objectives of the paper are to:

- review the literature on measuring health outcomes;
- consider the feasibility of measuring the inputs, outputs and outcomes of health nonprofits; and
- recommend what can be measured to determine the contribution of health nonprofits to Canadians' health and/or the health care system.

The paper begins with a working definition for the voluntary health sector for purposes of the research undertaken. A discussion follows as to the nature of the sector's value or contributions. Next, a set of generic activity domains that define the sector are combined with typical organizational performance models. This provides a conceptual framework illustrating how multiple kinds of value can be derived for the voluntary health sector. The paper then discusses briefly the main evaluation methodologies found in the literature and the current practices of measuring inputs, processes and outputs described in the literature and by key informants. Issues and limitations concerning types of outcomes and outcomes measurement has also been treated. The paper concludes with the key conceptual, measurement and ethical issues raised through the research, and recommendations focussed on the feasibility and utility of various measures of the voluntary health sector.

3 METHODOLOGY FOR THE LITERATURE REVIEW AND INFORMAL INTERVIEWS

The databases searched were ABI Inform, CBCA, Canadian Research Index, CINHL, Current Contents, Dissertation Abstracts, Econoline, Health Star, Medline, PsycInfo, Social Science Abstracts and Sociological Abstracts. Key words used and combined were: volunteer(s), voluntary sector, voluntarism, nonprofit organizations and benefits, measurement, evaluation, program evaluation, outcomes and impact.

The literature with respect to outcomes and measurement pertaining to the voluntary health sector in Canada was limited. This search was supplemented by a scan of the gray or unpublished literature in this area through an informal interview process beginning with the United Way of Canada. This source provided a list of other United Ways in the country that were known to be measuring or encouraging the measurement of program outcomes. Along with supplying information on what their organizations were doing regarding measurement, the United Way interviewees provided other contacts who were knowledgeable and experienced in outcomes measurement. Some were academics; others were private consultants with expertise in performance measurement; some were private citizens involved in voluntary organizations; and others were staff of voluntary health sector organizations. The key informant process was not intended to be comprehensive at this point in time since a sectoral consultation process has been viewed as a major input to this paper (The list of organizations contacted appears at the end of this paper).

The contacts were asked to elaborate on the following:

- their knowledge of how the voluntary sector measures itself;
- their knowledge of the health component of the sector if relevant to them;
- whether and how outcomes are measured;
- other contacts, and published and unpublished literature.

In general, the information gathered from all sources related more to the voluntary sector as a whole and less to its health component. Wherever appropriate, this paper presents the findings for the voluntary sector as generalizable to the voluntary health sector.

4 WHAT COMPRISES THE VOLUNTARY HEALTH SECTOR

The process of defining a voluntary health sector had two steps: the first identified the health organizations or organizations working in health within such a sector; the second identified which of these organizations' activities have a health outcome.

Beginning with the terms "voluntary" and "nonprofit," Hirshhorn (1997) and Rekart (1993) link the two by defining nonprofits as being prohibited from distributing profits to stakeholders and having a contribution of volunteer labour, at a minimum through volunteer boards of directors. Another source (Steering Committee, Social Accounting Framework Project, 1998, subsequently referred to as the Steering Committee) added that nonprofits have autonomy from the state and have democratic management.[2]

A sub-categorization presented by Day and Devlin (1997) and Reed (1997) makes a distinction between formal and informal activities within the nonprofit sector. Formal nonprofits are registered charities and non-charity (unregistered) organizations, and informal organizations and activities include local community groups or individuals, usually family and friends, who donate time and/or money for which no official record is necessarily kept.[3] While the significance of the informal sector is acknowledged, there is limited information available about it, for example, its level of activity, measurements of its contributions and even how to locate its constituents. For these reasons, the literature review and interviews were confined to the formal component of the voluntary sector. More research is needed on the informal component.

It is when introducing "health" as another qualifier after "voluntary" and "formal" that attempting to put boundaries around the voluntary health sector becomes illusive. Two organization-based classification systems specifically address and define health nonprofits: Revenue Canada's classification of registered charities whose purpose was stated as health on registration; and the International Classification of Nonprofit Organizations (ICNPO). The Revenue Canada category includes: hospitals, health services other than hospitals; health charitable corporations; health charitable trusts; and health organizations not classified elsewhere (Sharpe, 1994). For non-charity (unregistered) nonprofits, Revenue Canada has a number of categories, one of which is "other". The category contains several types of organizations including those classified as health-related.

Revenue Canada's classification of registered and unregistered organizations is broad. It spans a continuum of types of health services, from highly institutionalized organizations largely dependent on public funding, such as acute and long-term-care hospitals, to grass roots self-help groups that may rely exclusively on donations or members' fees for survival. Volunteer activities within these organizations can range from roles as community representatives on boards and/or committees, to involvement in fundraising and governance, to administrative support, to direct service to clients.

Compared to Revenue Canada, the ICNPO health category is less broad and is dominated by institutionalized health services. The grouping includes the following organizations: hospitals and rehabilitation facilities; nursing homes; mental health and crisis intervention services; and other health services (i.e., public health and wellness education, outpatient health treatment, rehabilitative medical services, and emergency medical services) (Canadian Centre for Philanthropy, 1999).

From a completely different perspective, the scope of voluntary sector organizations and activities with an impact on health can take into consideration the currently recognized definitions of health and the health determinants within the population health paradigm. Beginning with current definitions of health, they are holistic, with the World Health Organization (WHO) defining health as

> a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

Young (1998) supplies other definitions, one from the European Region of WHO that reformulated the WHO definition to align it with the population health paradigm as follows:

> [Health] is the extent to which an individual or group is able on the one hand to realize aspirations and satisfy needs, and, on the other hand, to change and cope with the environment. Health is therefore seen as a resource for everyday life, not the objective of living; it is a positive concept emphasizing social and personal resources as well as physical capacities;

and others that present an epidemiological point of view, where health is

> a state characterized by anatomical, physiological and psychological integrity, ability to perform personally valued family, work and community roles; ability to deal with physical, biological, psychological and social stress; a feeling of well-being; and freedom from the risk of disease and untimely death; or
>
> a state of equilibrium between humans and the physical, biologic and social environment, compatible with full functional activity.

Turning to population health, Hayes and Dunn (1998) found that the expression can refer to many kinds of activity. They cite the 1997 definition of the Federal/Provincial/Territorial Advisory Committee on Population Health:

> Population health refers to the health of a population as measured by health status indicators and as influenced by social, economic and physical environments, personal health practices, individual capacity and coping skills, human biology, early childhood development and health services.

As a framework, population health is activity-based as opposed to organizational. It relates major social, physical, behavioural and biological factors or activities to the overall health status between and within identifiable groups (Hayes and Dunn, 1998; Canadian Public Health Association, 1997). The broadness of the framework potentially captures a huge variety of activities and their associated organizations and assigns them a health impact. For example, the Nuffield Institute of Health describes health outcomes within the population health paradigm as

> the results (effects) on health of any type of process, for example, a health care service, housing, social service or employment.

The socioeconomic determinants of health are well documented within evolving research. For the purposes of our research, we felt compelled to limit the scope of the organizations and the activities with an impact on health to be included in a so-called voluntary health sector. Our decision had two bases: the first being the practical difficulties experienced by the field in relating program outcomes to population health described by a number of authors; the second being that data collected to date describing the voluntary health sector by, for example, Revenue Canada, have been limited to those organizations that recognized themselves as involved in health.

Elaborating on the practical difficulties, Kazandjian, Parnell and Lawthers (1995) state that even for health services, the correlation between what is done (process and output) and the results (outcomes) is often weak and unpredictable – one patient may survive grossly inappropriate or poor quality care while another may die even when everything was done right. Longo (1994) points this out while discussing the community health benefit of tax-exempt hospitals. He advises caution in interpreting community benefit outcomes because any specific outcome may be influenced by a variety of factors, only one of which may be under the control of the hospital. For example, health status, health behaviours, co-morbidity, age, genetics, education, income, culture, social class and ethnicity may all contribute to a particular health outcome.

From the broader perspective of human service agencies involved in activities that may relate to determinants of health, isolating the impact of a single intervention on health can be difficult, if not impossible, depending on the problem being addressed. For one, interventions need to be of a size and scale that the difference is detectable at the population level (Plantz, Greenway and Hendricks, 1997). Except in rare instances, individual programs do not serve enough people to affect community-level statistics. In addition, community-level conditions are

very often the result of a "constellation of influences" that are beyond the scope of a single program or intervention (Plantz, Greenway and Hendricks, 1997).

Researchers are also beginning to realize that the lead time for changes in population-level indicators can be longer than those demonstrated for individual-based interventions (Hancock et al., 1997). Where longer term outcomes measurement is undertaken, it also becomes difficult to determine causality between the outcomes and specific program activities (Plantz, Greenway and Hendricks, 1997; Webber, 1994) In other words, it is more difficult in the longer term to isolate that a particular intervention was responsible for a particular result because of other factors that may have had an influence on "the problem" over time.

As well, it seemed presumptuous to identify with a health sector those organizations that had not directly stated their purpose as health, regardless of the population health approach.

Therefore, as to which organizations to include and exclude from a voluntary health sector, we accepted the Revenue Canada health classification. It is more inclusive of organizations other than health institutions, and, in thinking ahead to a possible survey or case studies to gather information on the value of the sector, new data collected within the Revenue Canada framework would correspond to existing data and enhance the robustness of information on the sector.

4.1 Organizational Activity Domains

In order to refine the Revenue Canada classification and to add an activity-based dimension to it to organize and account for what the health-related organizations generally do, we developed a set of activity domains. The activity domains provide a means of relating discrete organizational processes to specific measurable health outcomes, and are sufficiently generic as to potentially apply to organizations working in areas other than health.

From the literature and from conversations with a number of people active in the nonprofit sector and in voluntary organizations working in health, five so-called organizational activity domains appeared to capture what voluntary health organizations do: they provide services (disease and illness prevention/health promotion and care); they do research; undertake advocacy; do fundraising; and regulate practitioners active in the field. Combining Revenue Canada's organizational classification with the activity domains produces a working classification for and definition of the voluntary health sector, as follows:

> The voluntary health sector contains formal organizations that are registered charities or nonprofit organizations without charitable status with a stated purpose or direct interest in health demonstrated through the following domains of

activity that contribute to health outcomes – service provision as disease and illness prevention/health promotion, and care; research; advocacy; fundraising; and practitioner regulation.

To the organizational activity domains we added a societal domain of contribution, recognizing that products and services of nonprofits are returned to the community. This is discussed in more detail in the following section.

5 THE SPECTRUM OF "VALUE" OF THE VOLUNTARY HEALTH SECTOR: A CONCEPTUAL FRAMEWORK

Webster's New World Dictionary states that value can be

> that quality of a thing...being more or less desirable, useful, estimable, important...that which is desirable or worthy of esteem for its own sake...

It could be assumed that value is synonymous with contribution. Measuring contribution first requires dissection of the sector into its component parts. The National Survey on Giving, Volunteering and Participating (Hall et al., 1998) produced information on the value of the voluntary sector based on a household survey targeted at individuals. The dissection proposed here, stemming from the Revenue Canada classification and evident in the conceptual framework, is organization-based. Second, determining value requires a further breakdown of organizational activities and processes because value is multifaceted. It potentially emanates in many forms (some concrete and some more abstract) from the activities or programs undertaken by an organization when it provides a good or service within its mandate. The value can be in economic and social (non-economic) terms. For example, the value of a good or service can express the power of purchasing other goods which the possession of the good or service conveys, or it can represent the utility of some particular object or service (Richmond, 1999). Value may or may not be measurable; it may or may not be additive with that of other organizations – the whole may be greater than the sum of the parts due to symbiotic relationships between organizations and programs.

The traditional model that demonstrates how an organization provides a good or service is illustrated in Figure 1.

Figure 1: Inputs through Outputs

Inputs → Activities → Outputs

Organizational inputs are the resources dedicated to or consumed in the process of providing a good or service that flows out of the organization. Typical inputs in a voluntary organization are revenues, assets, paid staff, their salaries and benefits, volunteers and volunteer time.

Since service delivery is linked directly to consumption in nonprofit organizations, that is, a service cannot be delivered unless someone receives it, inputs can also refer to demand inputs, which are expressions of need and client requests for service (Elkin and Molitor, 1984).

Activities or processes are what an organization does with the inputs to fulfill its mission. Put another way, it is how resources are arranged in terms of the goods or services needed to address demands (Elkin and Molitor, 1984).

Outputs or throughputs are the direct products of program activities and are usually expressed in units of service, for example, the numbers of clients served through the particular programs offered by an organization (Elkin and Molitor, 1984).

Plantz, Greenway and Hendricks (1997) expanded the traditional model into a chain of influences, with organizational inputs leading to activities, then outputs, and finally to a series of outcomes (Figure 2). Osborne and Tricker (1996) refer to this same chain of influences as the different parts of a service production process that make up a conceptual framework for studying organizational performance.

Outcomes are

> the benefits for participants during and after program activities (Plantz, Greenway and Hendricks, 1997).

A health outcome is

> a change in the health of an individual, a group of people or population that is attributable to an intervention or series of interventions

(Rissel, Ward and Sainsbury, 1996).

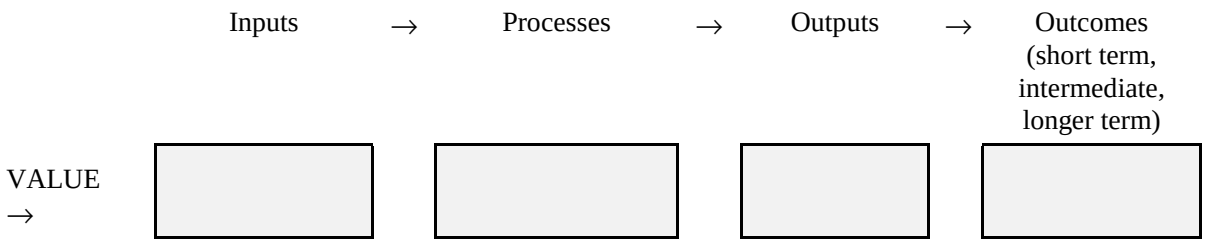
Outcomes can be confined to a short or intermediate term, or can apply to the longer term. The term "bridging goals" can be used instead of intermediate outcomes, these being outcomes that connect activities to ultimate goals (Thomas, 1994). The term ultimate outcome is sometimes used, meaning the final impact being sought to correct "the problem" with the provision of a good or service. Another notable term is proxy outcome, where a midway result is taken as an indicator that an ultimate outcome has been achieved, for example, using the falling incidence of a sexually transmitted disease like gonorrhea as an indicator that HIV transmission rates have also decreased.

The inputs through outcomes model in Figure 2 is the basis on which a discussion of the voluntary health sector's value can begin. Not only do inputs through outcomes define the framework through which organizational performance can be measured, they also represent the spectrum across which value can be derived. As Figure 3 shows, there is inherent value that can be measured at each point in the spectrum.

Figure 2: Inputs through Outcomes

Inputs → Activities → Outputs → Initial Outcomes → Intermediate Outcomes → Longer-term Outcomes

Figure 3: The Spectrum of Value: Inputs through Outcomes



Combining the organizational activity domains with inputs through outcomes as in Figure 3, was a first step in developing a conceptual framework for our research. It set in motion our thinking on outcomes and measurement of the voluntary health sector, as well as on what the sector contributes. (See Table 1. The components in the table are illustrative rather than exhaustive.)

Recognizing that the products and services of nonprofits are returned to the community, we added a societal domain to the organizational activity domains. Nonprofits' products and services add value that is more socially than privately distributed (Steering Committee,1998). In this way, the conceptual framework for our research respects the standard model of how an organization can measure performance at the same time that it advances the discussion of the kinds of value or impact that can be attributed to the activities undertaken. For example, from the perspective of their processes, all voluntary nonprofits can make a significant contribution within what we have called a societal domain. Here the Volunteer Canada (1998) adds that voluntary activity plays the role of a glue in Canadian society, holding all spheres of society together. Volunteerism can articulate unmet needs in the community and the organization of services to meet them. Voluntary activity celebrates the diversity between individuals and between social and ethnic groups (Knapp, Koutsogeorgopoulou and Smith, 1996). People seeking to reify

benefits of volunteering accrue not only to the organization or the individual but to the whole of society. For one, an agency relying on volunteers may provide training that a person could apply to other than their volunteer work for the benefit of other individuals or organizations. Second, the process of volunteering has benefits within the societal domain. For example, nonprofits that engage volunteers and provide services and caring are contributing to individual well-being and are enriching the social capital and cohesion of the communities in which the volunteers are active. In addition, nonprofits are described as: providing mechanisms of communication and consent for democratic social order; sustaining a sense of community and mutual responsibility; acting as political and social stabilizers by filling the service voids that governments and for-profits cannot fill; and influencing economic growth (Reed, 1997; Mulgan and Landry, 1995; Rose-Ackerman, 1996; Hirshhorn, 1997; Ross, 1994; Rifkin, 1995; and Drucker, 1989).

their ideas on how best to do something are more likely to find nonprofits in a position to serve that reification than either governments or for-profits (Rose-Ackerman, 1996). In this latter context, volunteering provides meaning more to the individual than to a vision of community (Van Til, 1988).

While the contributions to the societal domain are of unquestionable value, their measurement may pose a challenge.

5.1 Social Capital and Social Cohesion

Nonprofits exist in the "space between the individual and the state" and can positively influence social capital – the "density and quality of ties among persons and households"; the "features of social organization, such as networks, norms, and trust"; the interpersonal trust, reciprocity and collaborative behaviour in the pursuit of goals that are not strictly private and personal but rather serve some common good (Jensen, 1998; Veenstra, 1995). Nonprofits engage people in what Robert Putnam called "associational life" (Harriss and de Renzio, 1997).

The operationalization of the term social capital came about through the work of Putnum (1993) who argues that trust, reciprocity and civicness generate social capital that in turn impact positively on the economy and government. Further, he suggests that much of the social capital in a community is generated by local groups such as "singing groups and soccer clubs". In a subsequent study (Putnam, 1995), he reports on a decreasing stock of social capital in the United States.

Maxwell (1996) suggests that "social cohesion involves building shared values and communities of interpretation, reducing disparities in wealth and income, and generally enabling people to have a sense that they are engaged in a common enterprise, facing shared challenges, and that they are members of the same community". Jenson (1998) presents 5 dimensions of social cohesion: belonging and isolation; inclusion and exclusion; participation and non-involvement; recognition and rejection; and legitimacy and illegitimacy. In many ways, social cohesion appears to be a more robust description of social capital.

A key question to consider in examining the VHS is whether social capital and social cohesion are appropriate or, more importantly, feasible outcome/impact measures. In other words, can we measure them and is it meaningful? Jenson (1998) states that "social cohesion remains a contested concept". Lomas and Veenstra (1996) question the usefulness of Putman's measures of social capital because the underlying mechanisms that link social capital to good governance and well-being are not uncovered. Their exploration of these mechanisms may be helpful to measuring impact of the VHS in social capital terms.

Another key question relates to whether social capital is fostered by community organizations which, in turn impacts on individual traits such as trust and reciprocity. Jenson (1998) notes that Lomas and Veenstra suggest that "it is more helpful to look at the actions of institutions such as governments or the voluntary sector to understand community health". Indeed, some argue that social cohesion from the perspective of the individual is a determinant of well-being within the population health paradigm – people who are socially integrated live longer (Kawachi and Kennedy, 1997).

Table 1: A Conceptual Framework and Classification for Activities and Value of the Voluntary Health Sector

	ORGANIZATIONAL ACTIVITY DOMAINS						SOCIETAL DOMAIN
	Fundraising	**Research**	**Service**		**Advocacy**	**Regulation**	
			Prevention/ health promotion/ education	**Care**			
Inputs – resources	- volunteer hours; paid staff hours; $ resources						
Inputs – demand	- unmet financial organizational needs	- disease/illness prevalence	- risk populations - incidence of risk behaviours	- disease / illness prevalence	- unmet needs - quality of life issues	- accountability for practices, licensing, certification	
Processes	- events - materials - canvassing	- projects	- identification of populations at risk - programs - interventions	- programs - services	- lobby - media campaigns - networking	- regulation/ standards development - license and certificate issuance and revoking	- volunteer training - civic participation
Outputs	- # donors - $ raised	- # publications - # citations - # presentations - findings	- # materials published - # information campaigns	- # clients/ patients served - avoidable incidents	- # policy people contacted - # communiques	- # licenses issued & revoked - # professional practice complaints/ investigations	
Outcomes	- diversity of donor base - sustainability of donor base - expansion of donor base	- uptake and application of findings by practitioners and public - knowledge transfer	- % target population reached - changes in knowledge, attitudes and behaviour - decreased disease incidence	- maintenance of current health state - reduced rate of decline - cure	- policy changes - issues acknowledged or given profile - improvements in quality of life - services to meet needs	- public trust in a health profession - limited / protected access to practice	- development of social capital - contribution to social cohesion

Given the dearth of information on the VHS which includes the lack of a typology and basic economic information, concepts such as social capital and social cohesion and their measurement, while important and intellectually appealing, are most likely not first order priorities. Rather, our current knowledge of the sector would be best served by a “forward to basics” approach involving the collection of simple economic and non-economic data.

6 MEASUREMENT IN THE VOLUNTARY SECTOR

In seeking to determine the value of the nonprofit health sector, and having described the multifaceted nature of organizational and program performance to which value can potentially be assigned, a brief description of the performance measurement methods that evaluation science makes available to nonprofits helps us to understand what nonprofits may potentially know of their operations. The discussion here also raises our awareness of what the field is experiencing in terms of general measurement and outcomes measurement in particular.

6.1 Definitions

The literature search found numerous terms used in the contexts of measurement and evaluation. The following definitions were helpful.

A **goal** is a general statement of intent or aspiration (Rissel, Ward and Sainsbury, 1996).

Effectiveness means having an impact or effect (*Concise Oxford Dictionary*, 1982).

Program evaluation can be a comprehensive exercise focussing on four main areas of a program: rationale; processes; outcomes; and efficiency (Thompson, 1992).

An **indicator** is a specific measure for assessing progress toward a standard or benchmark for a rationale, a process or an outcome, or toward a level of efficiency.

6.2 Evaluation Models

Many classification systems are applied to evaluation science as it evolves. Scriven (1996) refers to one classification approach dealing with the formative and summative roles of evaluation. Formative approaches are designed, done and intended to support the process of improvement, and are normally commissioned or done by, and delivered to someone who can make improvements (Patton, 1996). They are typically time limited, meant to set the platform for subsequent summative evaluation (Patton, 1996). Summative evaluations are done for, or by, any observers or decision makers who need evaluation conclusions on which to judge a program and take action (Patton, 1996). Other evaluation classifications relate to methodologies such as quantitative versus qualitative; distanced versus interactive; and evaluator or collaboratively authored (Scriven, 1996).

Thompson (1992) attempts to organize the various evaluation models into four main paradigms. The first considers evaluation to be synonymous with applied research involving rigorously designed comparative studies; true field experiments; randomized controlled clinical trials; and quasi-experiments where experiments are impossible to implement. The evaluation is primarily summative, comparative and quantitative. Despite the scientific rigour, Rossi and Freeman (1993) write that these evaluations can only produce estimates with varying degrees of plausibility. In general, the more rigorous the evaluation design, the more plausible the resulting estimate.

The second paradigm locates evaluation within systems management. Outcomes are evaluated according to stated (given) goals and the primary audience is internal management, and program developers more specifically. These evaluations are formative.

The third paradigm defers evaluation to professional judgement, letting those with the most expertise make judgements on the quality of a program. Peers assumed to be objective provide reliable and valid assessments of a program by comparing it to peer-established standards. An example is accreditation.

Thompson's final paradigm links evaluation to politics. Evaluation-as-politics takes into consideration that even the most rigorous evaluation studies applied to programs with clearly defined goals may still have little impact on decision makers. In other words, decision making regarding policy development or resource allocation is not always a linear or rational process. In response, this paradigm has evaluation research studies directed at all major stakeholders who play a role in maintaining, modifying or eliminating a program. It sees a program as having stakeholders who compete with each other for a greater share of authority over program resources. Scientific quality is not the principle

This shift in the nonprofit evaluation literature has recently been studied and documented. Effectiveness evaluation in the nonprofit sector is a key issue vis a vis deriving value and effectiveness by definition is a key link to outcome(s) determination. Forbes (1998) has published a comprehensive review of effectiveness evaluations in nonprofit organizations over a twenty year period and notes that "an emergent or social constructionist approach" has surfaced in empirical studies over the past two years. With the emergent approach, "assessments of effectiveness are not regarded as objective facts but neither are they regarded as arbitrary or irrelevant. Rather, the emergent approach holds that definitions and assessments of effectiveness have meaning but that the meaning is (a) created by the individual or organizational actors involved, (b) specific to the context in which it was created, and (c) capable of evolving as the actors continue to interact" (Forbes, 1998, p.195). Hence, the interaction

standard in these evaluations. Rather they emphasize descriptive and qualitative data and are committed to addressing the multiple perspectives of the program held by the various stakeholders.

Patton (1996) proposes three other approaches outside Scriven's and Thompson's classifications and paradigms. He speaks of knowledge-generating evaluations, where findings do not lead to judgements of program merit or worth but instead to an increase in knowledge about a program's effectiveness for conceptual use – there is no change necessarily in what is done in the program but rather in how the program is thought of; developmental evaluations, where continuous program improvement or simply adjustment and adaptation are achieved with ongoing evaluative work (particularly suited for development programs where the process of engaging participants in setting and achieving their own goals is also a program outcome); and using the evaluation process to engage staff to think about and become more effective in their program efforts. The latter two evaluation designs have the purpose of directly enhancing a program's impact through the process of participant and staff engagement rather than with the use of program findings.

between the organizational actors and the environment that they function within creates an organic and evolving evaluation process - i.e. there is informal iteration. Evaluation becomes a process and not a discrete analytical objective.

The factors giving rise to the emergent approach are not well understood and it has been suggested that more needs to be known concerning external issues such as government cutbacks, tightened foundation requirements and intensified competition for revenues (Forbes, 1998). While at first glance the relevance of the emergent approach to empirical research on assessing the value and size of the voluntary health sector in Canada may appear weak, it may be beneficial to explore how effectiveness measures change over time, better understand the interactive effects and study interorganizational issues in order to derive fair, meaningful, and valid measures.

What voluntary nonprofit organizations share with any other organization intending to evaluate its programs is also critical to the debate over whether to choose qualitative or quantitative methods. Complicating the debate is that the quantitative and qualitative frameworks are not uniformly applied across the evaluation field, and both approaches continue to evolve (Hedrick, 1994). There is even argument over the definitions for the two methodologies. Reichardt and Rallis (1994) differentiate the two on the basis of paradigm, design and method of data collection. They describe the quantitative as operating within a "positivist" paradigm, rooted in scientific method that assumes an objective social reality that can be observed and measured. By comparison, the qualitative approaches assume a constructionist view where people subjectively construct their perceptions with there being no objective reality (Reichardt and Rallis, 1994; Waysman and Savaya, 1997). Reichardt and Rallis (1994) say that quantitative designs are experimental or quasi-experimental, involving representative samples and control groups where necessary to determine causality. The qualitative designs, on the other hand, base causal conclusions on the congruence of perceptions expressed during consensus-building processes. Regarding data collection, Reichardt and Rallis say quantitative methods are generally focussed on obtaining specific items of information using systematic approaches isolated to certain people or places while the qualitative are more likely to involve techniques that add perspectives, raise additional issues and accumulate detail with less attention paid to inconsistencies in data collection procedures.[4]

6.3 Practice

The literature and interviewees concur that the mainstay of data recording and reporting by nonprofits has been in terms of inputs, processes and outputs. Plantz, Greenway and Hendricks (1997) trace the aspects of performance that nonprofits in the United States have been measuring for at least 25 years. The authors say that the scope of performance measurement has expanded and has followed approximately the course below:

- financial accountability (the first focus);
- program products or outputs;
- adherence to standards of quality in service delivery;

- participant-related measures such as demographics and information about clients' problems or status prior to service;
- key performance indicators that are largely ratios critical to organizational functioning that are derived from various categories of inputs, activities, outputs and total costs;
- and more recently, client satisfaction.

For the most part, these measures have shown how much effort has been generated for how many people, but reveal nothing about whether the effort has made any difference – is anyone better off as a result of the services.

6.3.1 Inputs

In the literature, the most common valuation of the nonprofit sector as a whole relates to its inputs and is presented in economic terms. In these cases, inputs can be revenues, assets, paid staff, salaries and benefits, facilities, volunteers and volunteer time.

The economic value of volunteer services for Canada was first approximated for 1971 by Hawrylyshyn, published in 1975 by Statistics Canada. Economic volunteer services were defined as:

> Those activities which were done by a person outside the market but may have been accomplished by hiring a third-person from the economic market; they are distinguished from household work by the fact that the benefits accrue to someone rather than the volunteer or the volunteer's immediate family (Hawrylyshyn, 1978).

In 1978, Hawrylyshyn applied the concept of market replacement cost to again estimate the economic value of volunteer work. He first developed volunteer work categories and mapped them to market equivalents, assuming that volunteer labour was as efficient as paid labour. He then estimated the hours per week of work by type of work from a sample of volunteer request forms from the Metropolitan Toronto Volunteer Service for the period 1972 to 1975. With the average number of hours per week, the number of people requested to volunteer and the market equivalents in wage rates, he calculated minimum, maximum and mean estimates of the economic value per volunteer. Combining this data with a 1971 estimate of the number of volunteers involved in economic voluntary services, he calculated the total value of volunteer services in Canada for 1971 to be about $1,045 million or 1.1% of GNP.

Ross (1990; 1994) used the estimate of volunteer time provided to voluntary agencies from the 1986/87 Statistics Canada Survey of Volunteer Activity and average service sector wages to determine that volunteer work was worth $12 billion in 1986/87, excluding paid staff of the voluntary organizations and the value of informal volunteer activity. In 1993, it was worth $15.9 billion. He estimated that the hours volunteered in 1986/87 were equivalent to 617,000 full-time jobs.

A 1993 survey conducted by the CCP estimated that the entire charitable sector had $86.5 billion in revenues in 1993 ($90.5 billion in 1994); $82.4 billion in expenditures; $109 billion in assets; and constituted 1/8 of the GDP (Sharpe, 1994). Hospitals accounted for 30.4% of total charitable sector revenues, and other charitable organizations working in health accounted for another 5.8%.

The CCP survey also found that Canadian registered charities employed about 1.3 million people in 1993 – 9% of the total Canadian labour force – and were paid $40 billion in salaries and benefits: 468,000 of these (35%) were employed in hospitals and accounted for 41% of the salary costs. Other charitable organizations, including those working in health, accounted for 465,000 full and part-time people, with welfare and health organizations providing more than half of this employment.

Regarding volunteers, the CCP survey found that 107,000 people provided unpaid work in hospitals and another 600,000 contributed to other charitable organizations working in health. A 1997 Statistics Canada survey – the National Survey of Giving, Volunteering and Participating (NSGVP) – found that volunteering in health organizations (according to the ICNPO category) accounted for 10% of the total 1.1 billion volunteered hours in the 12-month period ending October 1997 (Hall et al., 1998).[5]

There are several methods available to value unpaid work. Two approaches consider the value of the time inputs (volunteer hours) using opportunity cost or replacement cost, or the value of outputs, that is, the market price for goods or services produced (Ecumenical Coalition for Economic Justice, 1997). Using the standard economic notion of value being based on willingness to pay, Brown (1999) argues that volunteering produces value according to at least two constituencies – volunteers and recipients. The value of volunteer services can be different from these two perspectives. For example, the value to volunteers (which can otherwise be called the value of their input hours) can be measured by what they willingly give up to volunteer (opportunity costs). To derive the value of what volunteers give up, Brown proposes refinements to the standard average wage rate, for example, using the after-tax net wage rate, accounting for the loss of fringe benefits with the unpaid work compared to paid work, and accounting for the stresses associated with paid work. In this way, Brown calculates that volunteer time can be valued at roughly 50 to 85 percent of the average hourly paid rate of the volunteer. Regarding the value of volunteer services according to recipients, Brown (1999) uses the standard economics approach of determining the clients' willingness to pay had they been given the choice between the volunteered services and cash. Brown concludes with three estimates for the value of formal volunteer labour in the United States for 1996 with the base being the average hourly wage for civilian, non-agricultural and non-supervisorial work – at least $113 billion based on recipients' willingness to pay; $203 billion acknowledging the approach that says that a part of people's paid wages compensates them for job stresses; and $317 billion using after-tax economy-wide hourly compensation rates.

A common valuation of inputs to the sector refers specifically to donations other than time, such as cash, gifts in-kind or indirect financial support. Cash donations are made up of financial donations to organizations, contributions to cash boxes and bequests; in-kind donations are clothing or household items and food; indirect support is given through the purchase of fundraising goods or services, or through charity-sponsored gaming (Hall et al., 1998). Data on the value of donations are limited to those for which receipts were issued by registered charities. For the voluntary health sector, Browne and Landry (1996) indicated that $682,488,000 worth of "gifts" were received

in 1993, representing 13.8% of the total value of gifts received by registered charities. Hall et al. (1998) found that the amount of donations in 1997 to health organizations was about $748 million, representing 17% of the total value of donations to charitable and nonprofit organizations.

6.3.2 Activities or Processes

Interviewees stated that nonprofits typically measure how they use their inputs in terms of the numbers and types of activities, programs and services they choose to deliver. They essentially describe an inventory of what is provided and in what quantity as the response to demand inputs, mentioned above.

It is the process of volunteering that brings to light several other aspects of value or contribution of the voluntary sector, some of which were mentioned in the previous section on the Spectrum of Value. One such contribution that is potentially measurable is skills development. The Points of Light Foundation survey (1996) found that voluntary health agencies in the United States utilize volunteers predominantly in the traditional areas of governance and fundraising and, to a lesser extent, in direct service. Volunteer Canada (1998) states that voluntary organizations have identified six general areas of skills training that they provide to volunteers: interpersonal; communication; organizational and managerial; fundraising; technical; and enhanced knowledge of the issues specific to the voluntary organization. Browne and Landry's (1996) data on staff and volunteer training by non-governmental groups and agencies (NGGA) comprising nonprofit corporations, charities, cooperatives and unions indicated that 34% of the health NGGA surveyed trained their volunteers. Thirty percent of these received top management training; 23% received other management training; 80% received service provision training; and 14% had clerical, administrative and other training.

Given this reality, volunteer experience has become an asset for those searching for paid work. Employers are increasingly accepting volunteer work as a valid part of work history and often look for volunteer experience as an indicator of initiative and a demonstration of potential and willingness to take new responsibilities. Young people particularly, who face significant pressure/obstacles in the transition from school to work, can gain hands-on experience in real work situations as volunteers with the skills, knowledge and technical experiences gained being

transferable to the labour market. For example, Hall et al. (1998) found through the NSGVP survey that approximately 24% of individuals aged 15 to 24 who had taken part in volunteer activities stated that these activities helped them obtain employment. Andrews (1990) found a strong relationship between volunteering and the psychological state of volunteers, including at least volunteers' perceptions that their physical health was superior to that of peers who did not volunteer. Andrews' study suggests that volunteer activity, particularly when performed frequently and directed toward strangers, is on the list of lifestyle and social network factors associated with physical and mental well-being.

6.3.3 Outputs

Nonprofits have historically measured the numbers of outputs or throughputs delivered as the direct products of program activities (Plantz, Greenway and Hendricks, 1997). The combination of inputs and outputs has provided many useful indicators for organizational management and has been the mainstay of data collected by funders.

One of the most advanced components of the voluntary health sector in terms of measuring outputs along with inputs and processes is hospitals. Since the early 1950s when the federal government began transferring funds to the provinces and territories to support aspects of what eventually became universal health care, hospitals have been recording and reporting inputs, activities and outputs in increasing detail and degree of standardization. The recording of inputs and particularly financial ones was dictated by first, the Canadian Hospital Accounting Manual originally published in the 1950s, and replaced by the Guidelines for Management Information Systems for Canadian Health Care Facilities in the early 1980s. The Guidelines give direction for recording financial and statistical data representing inputs, processes and outputs within specific functional frameworks that are intended to improve comparability of data among health care institutions. Most recently, data standardization may be seen as lapsing as provinces restructure health services in favour of relatively more autonomous regional health authorities, and provinces themselves become more autonomous in terms of how the envelopes of federal transfers are applied.

The early emphasis on data collection and standardization created an extensive body of data about hospitals compared to other less institutionalized health services. Typical output data for institutions relate to quality of care, resource management and service or patient throughputs. For example, acute and long-term-care hospitals track in-hospital infection rates; costs per case or per day; and average length of stay. The data are held by the provinces and territories and nationally at the Canadian Institute for Health Information (CIHI), Statistics Canada and Health Canada.

A number of interviewees indicated that other less institutionalized health sector organizations have also been collecting output data to meet their own management needs as well as those of their funders. The degree of standardization across these organizations would likely depend on the similarities in program focus and population, and on the reporting requirements of

Forty-six percent of this age group also believed that their volunteer activities had given them new skills that could be applied directly to their paid jobs or businesses.

different funders. Typical output measures relate to numbers of clients seen or served; number of contacts made; or number of products provided.

6.4 Outcomes and Outcomes Measurement: General

Outcomes measurement is gathering information related to anticipated results or changes to determine if these did indeed occur (Canadian Hemophilia Society, 1993). Rossi and Freeman (1993) write that outcomes measurement is to produce an estimate of the "net effects" of an intervention, where the estimated impact is uncontaminated by the influence of other factors that may affect the outcome at which the program under evaluation is directed.

Defining and measuring outcomes can be complex because outcomes have multiple dimensions. For example, the combination of possible measurement methods and the time frames for measurement can create a number of different outcomes for any activity being evaluated. This is illustrated in Diagram 1, where outcomes have three dimensions, placed along three axes. Along the *x* axis are the organizational and societal activity domains. Along the *y* axis are direct and indirect outcomes, direct being those that are the ultimate intention of an activity and indirect being proxies for the ultimate outcomes or bridges to them. The *z* axis represents outcomes determined by time.

Diagram 1: The Dimensions of Outcomes

z axis

Long term

Intermediate term

Short Term

	Organizational Activity Domain					**Societal Domain**	***x axis***
	Fundraising	**Research**	**Service**	**Advocacy**	**Regulation**		
OUTCOMES **Direct**							
Indirect							

y axis

Legend
x axis represents organizational and societal activity domains.
y axis represents the nature of outcomes.
z axis represents the timeframe in which outcomes may be measured.

The time periods range from the short term to the intermediate to the long term.

Diagram 2 represents the sequencing of an applied research project superimposed onto the input-output-outcome model and serves as an example of how a project flows from start to finish, resulting in so-called paybacks or positive outcomes (Buxton and Hanney, 1996). The outcome dimensions from Diagram 1 could be applied to the research project to expand its payback possibilities.

Plantz, Greenway and Hendricks (1997) write of the recent developments in the US nonprofit sector related to the uptake of outcomes measurement: national outcomes, the United Way of America conducted a telephone survey in 1998 of 35 organizations to learn qualitatively the status of national health and human service organizations'organizations are conducting research on outcomes; national organizations are developing resources for local agencies; managed care companies are stressing
service results; accrediting bodies are

Diagram 2: Inputs through Outcomes Model for Assessing the Payback from Applied Research

	Stage 0 - determining input demands	**Research needs assessment**	**Research sequence**
	Interface	*Project specification, selection, commissioning*	
Research project	Stage 1	Resource Inputs	
	Stage 2	Processes	
	Stage 3	Outputs	
	Interface	*Dissemination*	
Paybacks	Stage 4	Short-Term Outcomes	
	Stage 5	Intermediate-Term Outcomes (e.g., application of research; knowledge transfer)	
	Stage 6	Final Outcomes	

Source: Buxton and Hanney, 1996.

considering outcome-related criteria; and local agencies are operating effective involvement in outcomes measurement. The survey was in response to an outcomes measurement kit created in 1996 by the United Way of America and made available to any nonprofit interested in the exercise. Examples of the survey findings were: 94% of the 35 organizations surveyed had presented the topic at national or regional conferences; slightly less than half were providing staff training and technical assistance; and approximately one-third were creating lists of outcomes and/or indicators.

Of the 18 national organizations reporting local agency involvement in outcomes measurement, 78% indicated that the majority of their locals were undertaking outcomes measurement. The survey did not ask whether outcome measures had been developed or were in use.

Key informants indicated that large nonprofit organizations are indeed investing in outcomes measurement. Input, process and output data are no longer sufficient from the point of view of both public and private funders, the public as donors and current and potential volunteers. Some organizations are in fact using positive outcomes as promotional tools to attract resources. A number of individual United Ways in Canada are advancing outcomes measurement within their own operations and with their members. Some are offering members training, workshops and measurement tools, most often the United Way of America outcomes measurement kit, the Renssalaerville Institute milestone model, or other tools developed by the United Ways themselves.[6] Others have outcomes measurement systems. In terms of specific activities related to measuring organized outcomes measurement and data collection pilot projects among clusters of like member organizations or with samples of members.

The health promotion field, with its association with population health, is very active in developing the theory and application of outcomes measurement relevant to its work. Being very vulnerable in the current environment of results-based management and resource allocation, health promotion evaluators are closely following and challenging outcomes measurement that is limited to scientific quantitative methods focussed only on ultimate impacts. A number of key informants spoke of the need for, and in fact, actual shift away from quantitative methods toward the qualitative on the grounds that health promotion programs are best seen in a constructionist rather than positivist light where the outcomes to measure are negotiated rather taken as absolutes. One researcher and academic proposed that the key to evaluation of health promotion and education interventions is the accumulation of knowledge about the underlying processes that contribute to effective interventions rather than their ultimate outcomes. These key processes are 1) how social norms change; 2) whether or not an intervention is implemented as it should be (fidelity); and 3) audience reach.

6.5 Health Outcomes

The early work in measuring health outcomes dates back to the contributions of Donabedian (1969;1972) where evaluation

criteria relating to process and structure were used as proxies to actual measures of quality of care and health outcomes. The basic premise was that, if specific structural and procedural criteria were met, then it could be assumed that good quality of care was achieved. These concepts are currently used in the discipline of evaluation, especially where actual outcome measures are elusive or methodologically difficult to capture.

In many cases clinical or biological outcome measures were readily available in evaluating effectiveness. For example, hypertension treatment effectiveness is measured in terms of reduction in blood pressure and diabetes in terms of blood sugar levels. However, two problems arise in terms of the broader application of these outcome measures. First, there is the inability to compare the health outcomes on a cost-outcome basis due to the heterogeneity of the outcome measures (i.e. blood pressure versus blood sugar). The second problem is that functional health status is not included. For example, blood pressure could be decreased but impotence could be a side-effect of the antihypertensive drug therapy. Hence, analysts emphasized the need to integrate functional health status measures with clinical outcomes. Wilkins and Adams (1983) demonstrated the effect of considering quality of life by comparing data from the 1950-51 Canadian Sickness Survey and the 1978-79 Canada Health Survey. The findings revealed that between 1951 and 1978 life expectancy in Canada rose by 4.5 years for men and 7.5 years women. However, they noted a gap between life expectancy and quality adjusted life expectancy and emphasized the need to explicitly measure the “healthfulness” of life.

In the early 1970's economists began to examine this issue within the context of the economic appraisal of health programs. Torrance et al. (1972) developed a utility maximization model for evaluating health programs. Measuring utilities or patient preferences was viewed as the key to the development of a general outcome measure and thus negated the need to measure specific utilities for specific studies.

Generic health outcome measurement emerged from this seminal work and eventually gave rise to such instruments as the Health Utilities Index - HUI (Torrance et al., 1992).

The HUI Mark 3, the most advanced version, consists of a generic health status classification system and a preference-based scoring system. The health status classification system consists of eight attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain), each with 5 or 6 levels describing functional health status. The HUI has been used for calculating Quality Adjusted Life Years (QALYs) and the HUI Mark 3 has been included in every major Canadian population health survey since 1990 (Furlong et al., 1998).

Generally speaking, improving health outcomes involves improving quality of life and increasing length of life. The work of Torrance and his colleagues has lead to the development of composite measures of health outcomes such as QALYs, Disability Adjusted Life Years (DALYs) and Healthy Year Equivalents (HYEs) (Furlong et al., 1998; Buckingham, 1993; Gafni, Birch & Mehrez, 1993). All of these generic measures integrate functional health status or quality of life with increases in longevity and provide a common unit measure.

6.5.1 Outcomes and Economic Evaluation

The economic appraisal tool box of the health economist essentially consists of four approaches: cost minimization analysis (CMA), cost benefit analysis (CBA), cost effectiveness analysis (CEA) and cost utility analysis (CUA). These evaluation techniques have been described and a comprehensive description of the elements of sound economic appraisal has been published by Drummond, Stoddard and Torrance (1987). CMA is employed when comparing two interventions/programs that have equal outcomes. In CBA, cost and outcomes are expressed in monetary units. CEA and CUA use costs in the numerator and an outcome measure in the denominator (e.g. cost/QALY). While some analysts consider CUA as a particular form of CEA, one key difference relates to the outcome measure. CEA generally employ a natural unit of measure such mortality or morbidity relative to costs (e.g. $ to prevent HIV infection) and CUA use utilities or the desirability of a specific outcome.

The approaches to economic evaluation and the guidelines produced by Drummond and his colleagues may be of value in considering methods for assessing impact and value in the VHS. Specifically, insight may be gained related to developing a schedule of costs, the perspective or viewpoint of the study and the appropriateness of specific outcome measures. With respect to the latter, it may be feasible to examine the number of QALYs generated per year by the VHS or on average by VHS organizations.

6.5.2 Macro-Oriented Measures and Indicators

In Canada, the Canadian Council on Health Services Accreditation (CCHSA) researched indicator initiatives taking place across Canada in 1997. It searched the literature and the Internet, and interviewed health care professionals and organizations identified as leaders in their fields. The findings relate to a range of nonprofit health care organizations and to specific services within them. One of the observations was that there is no consensus on the definitions of performance indicators and process or outcome measures. The most highly institutionalized organizations, such as hospitals, are taking a "corporate" approach to indicator development, using a balanced scorecard[7] or a dimensions-of-quality approach that addresses the whole operation and emphasizes utilization and risk. On the other hand, community-based services and regional systems are shifting their focus from indicators attached to individual clients to those that present the community and population health perspectives. Across the spectrum of services, the cost of information systems for collection and monitoring of indicator data is an issue.

Along with the balanced scorecard approach, institutions like hospitals monitor other aspects of administrative operations such as human resources, information management and environmental management. As noted earlier, clinical services such as medicine, surgery, obstetrics and gynaecology rely on indicators of clinical process (outputs or throughputs and their quality) more so than ultimate outcomes of interventions. The same can be said of cancer treatment services, where most indicators focus on organizational processes. At the clinical level, most outcome indicators relate to specific types of cancer or treatment, and evidence-based clinical practice guidelines are gaining prominence and are also

influencing a trend toward clinical outcomes.

On continuing care, CCHSA found that most work on performance indicators of any sort was being carried out in large long-term-care organizations, acute care institutions with geriatric beds and by governments. In home care, the provider/purchaser environment is driving the development of indicators that demonstrate adherence to process standards and are related to price.

Rehabilitation agencies are facing greater challenges in measuring outcomes by virtue of the nature of the services. There is a lack of consensus on which outcomes should be targeted for measurement and on how and what the various aspects of rehabilitation separately contribute to rehabilitation.

Mental health services specifically, and community health services in general, have advanced relatively further in terms of identifying the outcomes to which they relate. Both articulate the broader determinants and risk factors of health as what they are ultimately seeking to influence. Community health care programs in particular are using a population health promotion framework for identifying outcomes and measuring them. Nonetheless, both types of services are struggling with determining which endpoints and outcomes are realistic for them to appropriate and measure.

Overall, the advancement of outcomes measurement within the Canadian voluntary health sector – producing an estimate of the "net effects" of an intervention – appears spotty at this time. Informal interviews have produced anecdotal information suggesting that large and sophisticated nonprofits are applying outcomes measurement to strategically apply and attract resources, while many small organizations are struggling with how and whether to redirect limited resources to measure outcomes. Many parent or national organizations are encouraging members to measure outcomes by supplying training and assistance. However, in the absence of standardized measures and with the perceived high need for nonprofit sector accountability, outcomes measurement in general is developing in a reactive manner (Steering Committee, 1998).

With regard to collecting outcomes measurement information as a means to determining the value of a number of voluntary health organizations, the individual organizationally defined outcomes would each point to a contribution but would likely not be additive to give a "bottom-line" for the sector. In other words, outcomes measurement across a broad and diverse voluntary health sector is not standardized to the extent of having common units of measure.

7 ISSUES, LIMITATIONS AND OBSTACLES TO DETERMINING THE VALUE OF THE VOLUNTARY HEALTH SECTOR

As previously mentioned, the measurement of inputs, processes and outputs in economic and quantitative terms has been the historic mainstay of data recording and reporting for nonprofits. Currently, budget reductions have and continue to create a competitive climate in which nonprofits generally are being scrutinized and compared on the basis of fiscal and programmatic accountability. Under these circumstances, nonprofit outcomes measurement ideally requires three conditions: input from stakeholders; accurate measures of a nonprofit's results; and uniform measures across subsectors of nonprofits, with health agencies being an example (Steering Committee, 1998).

The conditions pose a significant challenge to outcomes measurement. The challenges generally relate to there being no standard measures of nonprofit benefits; inconsistencies in methods of record-keeping and reporting among nonprofits; and the absence of agreement on outcomes and their measurement (Richmond, 1999). These and other challenges are discussed below.

7.1 The Debates over What and How to Evaluate

As would be the case with any program, the method of choice for its evaluation is dependent on a number of factors. Apart from financial and knowledge capacity considerations, examples of the factors are the nature of the program (having relatively objective or subjective impacts), the purpose of and audience for the findings, whether processes or outcomes are of interest, etc. Rossi and Freeman (1993) write that although qualitative methods and data are important for certain evaluative purposes, precise assessments of program impact require carefully collected quantitative data that relate activities to ultimate outcomes. However, in the context of evaluating nonprofit programs, other writers caution that discounting qualitative data can too severely limit an assessment. The goals of nonprofits may be too diffuse or too broadly based to be measured only by quantitative techniques (Thomas, 1994; Abbott et al., 1997). The value of qualitative measures is in their potential to capture non-numerical in-depth descriptions and understandings of program operations as opposed to focussing only on ultimate outcomes. Thomas (1994) endorses evaluation that begins with a goal of balance rather than objectivity. Where objectivity implies taking an unbiased view of a program from a distance, balance implies viewing a program from up close as well as from afar, using both quantitative and qualitative data.

Selection of evaluation methods is further complicated for voluntary nonprofits by their having a wide and complex range of stakeholders (more so than for-profits), resulting in a multidirectional accountability framework, for example, downward to staff and volunteers, outward to external stakeholders such as beneficiaries, and upward to board members and funders (Voluntary Sector Roundtable, 1997). Nonprofits are accountable to those stakeholders who can exercise sanctions, that is, donors and funders. They have an additional special duty to demonstrate their accountability to stakeholders who cannot exercise sanctions, namely, the beneficiaries (Plummer, 1996). As well, nonprofits are accountable to the general public, their

partners where relevant, their members where relevant, their volunteers and staff. Given the variety of audiences to which an agency may be accountable and asked to demonstrate positive performance, the selection of what to measure and how can vary. While the range poses a challenge for measurement, it also serves to highlight the variety of perspectives from which measurement and value need to be understood.

Often the first and simplest outcome being measured is client satisfaction (Plantz, Greenway and Hendricks, 1997; Quality Standards Task Group, 1998). While a step in the right direction, the Quality Standards Task Group (1998) states that community or client satisfaction is seldom sufficient as a measure of quality performance. It cautions that the indicator has limited value, in that a voluntary organization may be in a monopoly position in the provision of a particular service or function suggesting that high demand levels might not be an indication of satisfaction. As well, beneficiaries or end-users often lack significant power to influence the quality of the services they receive while some end-users may simply not be able to respond to questions about quality and stay silent, for example, beneficiaries of advocacy or
One could expect that with a variety of stakeholders having an interest in a voluntary agency, each could have a different idea of the value of a program or service and their opinions and demands may change over time. Wholey (1996) writes that the first and most difficult step in performance measurement may be in reaching agreement on the set of performance dimensions that are important to capture and then the quantitative or qualitative indicators that best represent those dimensions. Forbes (1998) says that the criteria by which organizations are

advisory services, mentally ill people or the very elderly.

7.2 The Subjective Nature of Value

Determining what aspect of performance to measure can be a moving target. In other words, value has a constructionist tone. Newcomer (1997) says:

> [value or outcome] is not an objective reality to be measured... [it] is a socially constructed reality [that depends on the] criteria for defining and rating [that] must be assembled and agreed upon by a group of officials with both the authority and responsibility for the [programs to which value is to be assigned]... stakeholders with interests in... a program... may [choose and rank] potential criteria differently.

Forbes (1998) writes that the definitions and assessments of organizational performance have meaning but that meaning is 1) created by the individual or organizational actors involved; 2) specific to the context in which it is created; and 3) capable of evolving as the actors continue to interact.

determined to be effective are negotiable and malleable. Herman and Renz (1997) say much the same, having found that stakeholders frequently vary substantially in their judgements of organizational effectiveness and indicator selection.

7.3 The Reluctance to Measure Outcomes

Overall, nonprofits perceive the largest barriers to their conducting assessments of program outcomes to be the lack of financial support, skills among staff and technical assistance (Zacharakis-Jutz and Gajenayake, 1994; Rodriguez-Spagnolo, 1992; and Taylor and Sumariwalla, 1993). For health organizations, more research on improved measures and instruments is a strategy to increase evaluation, second to increased availability of funds (Taylor and Sumariwalla, 1993). The reluctance can be easily imagined as the 1993 CCP survey of the voluntary sector found that 60% of registered charities reported fewer than two paid staff and accounted for 3.4% of total salaries and benefits.

Plantz, Greenway and Hendricks (1997) mention the tension between the need for technically sound outcomes measurement methodologies, which can be expensive and time consuming, and the staffing, funding and workload realities that constrain almost all health and human service organizations. Taylor and Sumariwalla (1993) add that applying purely scientific or quantitative inquiry to measuring human behaviour and changes to it can be difficult because of a number of factors: the number and types of variables that add to the complexity of assessment and which are a problem to control; confidentiality of client information; many effects of human service programs take a long time to show results, requiring longitudinal study to conclusively determine Outcome measurement among nonprofits is being promoted and is on the increase (Panel on Accountability and Governance in the Voluntary Sector, 1999; subsequently referred to as the Panel). Both funders and the public are asking for greater accountability for resources consumed by the voluntary sector. In this context, the success or failure; limited agency capacity; and the issue of cost – only well-endowed groups can afford high-quality and, where necessary, longitudinal evaluation. In a constrained fiscal environment where organizations can understandably be competing against each other for funding, the well-off organizations with sophisticated outcomes measurement results may have a significant advantage in terms of attracting funds over smaller organizations unable to undertake comparable scales or quality of evaluations.

Small organizations unable to conduct their own evaluations can turn to evaluative research on programs equivalent to their own that demonstrate positive outcomes. However, McNamara (1999) suggests that some nonprofit leaders have been discouraged by evaluation research because methods were chosen on the basis of achieving rigorous scientific accuracy, reliability and validity. Conclusions were limited and generalizations were avoided because of the scientific orientation of the evaluation. Key informants agreed that the knowledge transfer from research to the front line has in some cases been limited. Organizations often simply cannot duplicate in their day-to-day operations the strict research contexts that produced positive outcomes.

7.4 The Influence of Funders

Panel (1999) acknowledges that the shift away from traditional measures such as inputs and outputs to outcomes has intuitive logic in the voluntary sector but that implementation is complex. The Panel sounds a note of caution to funders and governments in that they need to understand the limitations of outcomes measurement

before imposing it on the voluntary sector. The Panel cites the lessons to funders that Plantz, Greenway and Hendricks (1997) compiled where, if funding is inappropriately linked to outcomes, resources would shift from services to measurement with no direct benefit to programs; prevention and development programs and others with harder-to-measure outcomes would be potentially penalized; "creaming" would be promoted (the deliberate selection of programs or clients certain to have positive outcomes); risk taking and innovations would be discouraged, as would interprogram/interagency cooperation because of a more competitive funding environment. Organizations might also choose inappropriate indicators of impact in order to demonstrate some effectiveness, for example, public health units giving a high profile to vaccination rates when the threat of infectious diseases is limited.

Most governmental calls for outcome measures are to inform allocation decisions (Newcomer, 1997; Thomas, 1994). However, when only economic outcomes analysis is used in resource allocation, the incentives to "cream" are obvious and pose a significant ethical dilemma. Agencies can be torn between what is called a substantive principle of justice, that is, serving the neediest first and the need to demonstrate positive outcomes (Yeo, 1993). For nonprofits, the neediest people may be difficult to serve and, as a result, the
The implications of an emphasis on short-term outcomes are even more significant. When only limited resources for intervention are available, the most efficient allocation would direct resources at the least affected populations (Birch, 1997). The greatest intervention benefits would accrue to those whose underlying conditions were most closely associated with the orientation of the agencies risk poorer outcomes that may jeopardize funding.

In writing about measuring clinical outcomes in the context of health maintenance organizations in the United States, Mosser (1996) speaks of the strategizing around what programs and outcomes to measure to yield the greatest gains across the organizations. He suggests, for example, that nonprofits select programs that address high frequency health care conditions in the population served; have the probability of achieving improvement in the process of care for the condition; and anticipate a significant magnitude of improvement in health outcomes for the conditions or reductions in cost or both. A careful selection of populations to serve, that is, people with a higher probability of demonstrating a positive outcome, seems a fairly reasonable extrapolation of Mosser's suggestions.

An emphasis on economic analysis of short-term outcomes can ultimately be a dis-service to nonprofits because other measurement approaches with formative designs to improve program operations may be more useful (Thomas, 1994). Given the choice about the use of limited resources to develop performance measurement systems, collecting data that will support budgetary requests will likely win over collecting internally useful but less resource supportive data (Newcomer, 1997).

intervention. For example, anti-smoking campaigns have largely influenced people of higher socio-economic status and have been less effective with more marginalized populations such as the unemployed, single mothers and aboriginal populations. The underlying health determinants of the harder-to-reach groups that are influencing their smoking patterns are not being

addressed by the campaigns. With the latter group, the most useful application of program evaluation would be through an iterative process with empirical evidence on campaign effectiveness collected over a period of time from a series of interventions. If such a longer term investment allowing "course correction" of the program over time were forfeited in favour of programs that show positive outcomes in the short term, the evidence-based decision making and resource allocation would be leading to increases rather than decreases in the inequalities in health status.

Zacharakis-Jutz and Gajenayake found that the reluctance to evaluate is often based on a lack of control over how the evaluation is conducted; concern over how results will be interpreted and used; and fear of how the evaluation might affect funding. Key informants concurred and added that nonprofits are sometimes reluctant to measure outcomes because of a lack of direction from funders as to what outcomes they will recognize. Organizations are unwilling to invest limited resources or redirect program funds to measurement that is not guaranteed to benefit them. Nonprofits also sense that government policy development and funding is not necessarily linear or rational but more political, in which case objectively demonstrating effectiveness may not guarantee a positive response from funders.

Funders, on the other hand, are frustrated by a lack of standardization and comparability in the outcomes that nonprofits measure. Their frustration however is one side of a trade-off situation. The other side emphasizes the need for organizations to define outcomes that directly relate to their specific missions and goals and that can demonstrate their program nuances. With their outcomes defined and measured, organizations can then approach funders with congruent goals (Plantz, Greenway and Hendricks, 1997).

7.5 The Diffuse Nature of Some Voluntary Health Sector Contributions

Nonprofits face challenges to demonstrating their value or effectiveness that for-profits, for example, do not. For-profits have private sector tests such as profit or return-on-investment as measures of success while nonprofits have no such bottom lines (Quality Standards Task Group, 1998; Young, 1999). Furthermore, unlike for-profits, nonprofits cannot eliminate service to clients who represent the greatest cost and the least tangible results (Steering Committee, 1998). These may in fact be the very populations that society expects nonprofits to serve.

With challenging mandates, a mission statement becomes particularly important to nonprofits as it justifies allocations and focuses the organizational resources into activities and outcomes and acts as a surrogate bottom line (Bart, 1998). The advantage of having the organizational mission statement reflect outcomes is that each organization is free to represent itself according to its mandate. It can have an accountability framework that fits its unique role (Steering Committee, 1998). The disadvantage may be that there is no set of standardized and accepted bottom-line types of performance indicators for nonprofits comparable to the financial measures of for-profits. Comparisons between nonprofits can therefore be difficult. The absence of standard measures also leaves nonprofits describing their outcomes in their own terms, which may be labelled as "soft" or subjective (Steering Committee, 1998).

Hirshhorn (1997) proposes that by meeting specialized and sometimes controversial needs that governments and for-profits cannot address, nonprofits find themselves often providing complex social and health services whose outcomes may be inherently difficult to assess. Examples are health promotion and disease prevention programs, and community action interventions seeking legislation or policy change, initiatives that attempt to implement changes that will simultaneously affect many individuals. These are difficult to evaluate in rigorous scientific terms because they may require communities to be the control groups, which can be impossible to operationalize (Hancock et al., 1997).

Webber (1994) adds that the goals of health and human service organizations are multifaceted and that the value of these organizations lies in issues of quality of care that are difficult to quantify, particularly if the outcomes being sought are in terms of economic "income" or economic contribution to society. Most difficult is how to measure program effects in terms of concepts such as client productivity or resource consumption, that is, what client characteristics to measure as indicators. As well, if the economic benefit to society is to be matched with the expense of a program, Nonprofits are responding to pressure for accountability from three principle sources: governments and other funders; the nonprofits' own boards and volunteers; and the general public. They are gradually realigning their resources to collect the information they need to demonstrate effective performance to all these audiences, especially in terms of outcomes. Because there are few limitations on what nonprofits can disclose of their operations, any new information about performance will likely become public. In this context, there are programs that increase human capital are presumed to have very long-term benefits which cannot be observed in the same year as the financial expenditures of an organization. Also difficult is measuring what would have happened in the absence of the client's entry into a program. Before and after comparisons are not sufficient because even in the absence of intervention, changes can occur over time. Cost-benefit analysis – the type of evaluation that is most likely to illuminate gross societal income – typically requires control groups, which, because of time, expense and for ethical reasons, are often beyond the capacity of many health and human service organizations.

As discussed earlier regarding the perspectives from which value of the voluntary health sector can be derived, measurement of the sector's impact within the societal domain is a challenge. Social capital and social cohesion are concepts for which units of measurement may be difficult to derive even for experts in the field not to mention voluntary nonprofits with limited resources.

7.6 The Ethics of Information Disclosure

implications and ethical issues related to the increased exposure of these organizations. These relate to privacy, accuracy of data released, "creaming" and misinterpretation of data and information, explained below.[8]

Privacy

An organization may be inhibited from presenting certain information or data because exposure would contravene the value it places on the privacy of its clients or staff.

Accuracy of Data

Where there are no standards, benchmarks or audits of other than financial performance results (as would be the case with many individual nonprofit outcome measures), organizations have an opportunity to present themselves in the best possible light and may feel compelled to do so if funding is to some degree dependent on performance. In this case, organizations are writing their own "report cards" and the accuracy of the material reported can become suspect.

"Creaming"

As mentioned earlier, in order to appear as good performers or to score higher, organizations may change their behaviour by adjusting program intake to gather clients more likely to contribute to positive outcomes.

Misinterpretation of Data and Information

Data may be released without sufficient background or interpretation, leaving out significant nuances that can harm the organization represented.

Why information is being provided needs to be clearly understood. The consumer ethos and government's evidence-based decision making need to be sensitively examined because of the potential harmful consequences of exposing certain services or organizations. Exposure needs to be balanced with the other values of the organization, for example, an organization may resist exposure of certain information for the good of a service or some of its aspects.

Nonprofits may feel that they are being herded into a sophisticated system of performance measurement (particularly with respect to outcomes) and public disclosure of operations that is not required of for-profits or even government programs.[9] As well, measuring the various perspectives of value described earlier will also advance significantly the information available for analysis of the voluntary sector, for example, creating the potential to link the value of inputs to outcomes. This may exceed the public information available about the for-profit sector and create an imbalance in the degree of exposure and scrutiny of nonprofits compared to for-profits. The exposure may place nonprofits at a disadvantage in a situation where they are competing for service contracts, for example, in regionalized or integrated health care systems. In these cases, the information would be feeding a competitive environment while the health care system is supposedly striving for more collaboration.

8 KEY FINDINGS

The findings of our research will inform first a consultation with representatives of the voluntary health sector, and second, possibly a survey of the sector and the development of case studies as a means to deriving, or at least understanding, the sector's contributions not only to the health of Canadians but also to the broader social fabric. We assume that a survey and the case studies would target organizations rather than individuals to capture a collective organizational viewpoint.

The findings appear in four groupings: the types of published material found; the perspectives on the value of the voluntary health sector activities and contributions; the implications of the measurement methodologies available; and the obstacles and ethical issues that arise while considering the derivation of value for the sector. Key opinions and the interviewees' comments are interspersed throughout the groupings.

8.1 Information about Outcomes and Measurement in the Voluntary Health Sector

Beginning with general measurement, voluntary health sector organizations can fairly readily supply information about their inputs through outputs, most likely in quantitative terms. This paper proposes five organizational activity domains (as illustrated in Table 1) as a means of further grouping data within and across organizations. Examples of data are numbers of active volunteers and the number of hours they contribute per activity, such as fundraising or advocacy. As well, numbers of paid staff and their salaries and benefits are available, as are the variety of programs and services provided and the numbers of clients served or products distributed.

Apart from measures of inputs and outputs, there is very limited published material about what the Canadian voluntary health sector considers to be its value in terms of health outcomes and how this is measured. Interviewees confirmed that among nonprofits, experience with outcomes measurement is in its infancy. The literature on nonprofits tends to be directed toward how to manage nonprofits rather than to their larger collective role in the economy and how their work should/could be measured. There is likely a body of organizationally specific in-house (gray) literature on performance evaluation and program outcomes measurement that can be best accessed through consultation with representatives of the voluntary health sector.

Regarding the benefits of volunteer work, the majority of the literature presents the perspective of the individual in terms of his or her qualitative assessments of personal benefit, and only a few studies are conducted with a societal focus.

8.2 The Various Perspectives on the Value of the Voluntary Health Sector

The first key challenge of this research was to define the "value" of the voluntary health sector. Nonprofits have historically represented their value in positivist terms, related to a typical model of program or service delivery that presents in economic and otherwise quantitative units the inputs, throughputs and outputs of nonprofits, such

as paid staff, volunteer hours, numbers of programs offered and numbers of clients served. Relatively recently, a generally heightened sense of nonprofit sector accountability has added outcomes to the program model as meriting identification and measurement. It is by adding the measurement of outcomes or impacts on health that voluntary health organizations have the potential to demonstrate other dimensions of their contributions aside from the historic measures of inputs and outputs. Many nonprofits are now attempting to or are demonstrating how they make a difference to the clients that they serve.

The measurement of outcomes involves a more judgmental assessment of value given that degrees of success or failure are captured. For example, a needle exchange program is directed at reducing the incidence of HIV among injection drug users. Certain levels of increase in incidence would be evidence of failure. The difficulties in attributing success and failure discussed earlier would apply in this case. In addition, the effort required to derive and attribute these measures may be prohibitive for many non-profit organizations.

Nonprofits contribute collectively to what may be seen as abstract commodities such as social capital and social cohesion. While less concrete than other measures, the value that is reflected in these latter concepts is generally not disputed. The major problem, however, is in quantifying and otherwise assigning a value to them. Hence the practical feasibility of measuring certain outcomes will be a major consideration for On the other hand, a health services planner may prefer to evaluate the extent to which the members of the Aboriginal community contributed to the development of the culturally appropriate service, this being an indicator of community attachment to and

any sectoral study focussed on demonstrating impact, contribution or value. The Steering Committee (1998) experience is of particular interest in this context. The Committee oversaw the application of a social accounting method (or social audit) that measured the outcomes, impact and benefits to their communities of five community-based training sites in Ontario. Using evaluation criteria and a participatory evaluation approach appropriate to nonprofit culture, the method attributed a comparative economic value to nonprofit activities and highlighted their contributions. Specifically, it provided a social return-on-investment ratio as an indicator of social accountability.

Another key finding related to the value of the voluntary health sector is the notion that what is deemed valuable is a socially constructed phenomenon rather than an absolute. The key study by Forbes (1998) documents the emergent approach to non-profit effectiveness assessment that has surfaced in the literature over the past two to three years. What is of value is specific to a certain context and dependent on the personal values of the stakeholders whose task it is to determine what is worth measuring in a program or service. For example, an Aboriginal community may decide that the most valuable aspect of a voluntary health sector service is its availability in a culturally appropriate format (a feature of process) while the funder of the service may regard final outcomes as warranting the most attention.

support for the service. With a sectoral study or survey, decisions will have to be made as to which and/or whose perspectives on value will be privileged.

8.3 Measurement Issues

The question of how to measure value is linked to what is deemed of value to measure. Health sector nonprofits have historically measured organizational and program-related inputs through outputs, largely driven by the reporting requirements of their funders. These same funders are more recently placing greater emphasis on value-for-money decisions for resource allocation, which require program outcome information, usually in terms of cost-benefit. Such an emphasis by funders can narrow significantly the nonprofits' options as to what to measure because of their limited resources. It may, in fact, be a dis-service to nonprofits because they may forfeit other more useful evaluation approaches in order to satisfy funders. As in the example above with the Aboriginal community, process evaluation may be more useful to an organization than outcome evaluation. Key literature sources and interviewees stressed the importance of organizations identifying for themselves the outcomes or other aspects of program delivery worthy of measurement as opposed to measuring what is of interest to funders.

Program outcome evaluation can by itself be a complex undertaking and be so to such an extent as to discourage smaller nonprofits from attempting the exercise. On the other hand, evaluation science is delivering an increasing array of methodologies for measuring not only program outcomes but also various other aspects of the program delivery process in order to provide organizations with the most useful information possible to maintain optimal programming. A limited review of the evaluation science literature found a broad array of evaluation methods. They range from rigorously scientific and quantitative experiments using randomized trials and control groups to prove program impact, to evaluations that deliberately involve many participants and collect numerous viewpoints in order to attach as large a constituency as possible to a particular program. In the latter case, the engagement of a large number of participants in program evaluation has a collateral benefit to the program itself.

Nonprofits may also be faced with the possibility that some of their programs defy measurement in practical terms. For example, an impact evaluation of a health promotion program targeted at the general population may require a longitudinal study that an agency cannot afford. Furthermore, both the literature and several interviewees proposed that nonprofit programs may be more suited to qualitative and grounded approaches directed at processes as well as outcomes, as opposed to strictly final outcomes measurement using quantitative methods.

Regardless of the richness of evaluation science, the barriers to program evaluation by nonprofits are for the most part lack of funding, the need for more training and assistance, and the need for research-based comparative indicators. In some cases, organizations are reluctant to measure their performance for fear of negative repercussions from funders if results are poor, or for fear of a refusal by funders to recognize results that they do not value.

8.4 Ethical Issues

In the context of performance measurement, most of the ethical issues relevant to nonprofits relate to the release of performance measurement results. Key issues relate to violations of the privacy of nonprofits' clients and staff; the incentives

to “cream” clients; nonprofits being compelled to present only the best performance results to avoid criticism or penalty; and the chance of misinterpretation of information. As well, nonprofits competing with for-profits for contracts in integrated or regionalized health services systems may be disadvantaged by requirements to disclose information that are not applicable to for-profits.

9 CONCLUSIONS

The goal of this paper is to facilitate the analysis and measurement of the economic and social contributions of the voluntary health sector, with the latter ideally presented in both qualitative and quantitative terms. The objectives are to consider the feasibility of measuring the inputs, outputs and outcomes of health nonprofits and to make recommendations on what can be measured to determine the contributions of health nonprofits to Canadians' health and/or to the health care system. The data needed to derive value may be collected through survey and/or case studies.

This paper discusses various measures and measurement issues whose relevance to the larger project objectives (i.e., the size, scope and value of the voluntary health sector) boils down to feasibility relative to research design and methodology.

9.1 The Feasibility of Research Design and Methodology

Presumably, the unit of observation for the voluntary health sector study is the organization. The capacity of the average organization to produce all the data discussed is a critical feasibility issue (specifically inputs through outcomes, as well as any contributions to the societal domain). Within the context of a survey of a cross-section of organizations, some of the larger organizations may have the required information "on board," whereas smaller organizations may have to devote resources to temporary data collection or make "educated guesses."

If it is assumed that the organization is the unit of observation, then it follows that the geographic universe must next be defined.

Since it has been demonstrated that value can be extracted discretely from inputs, outputs and outcomes, and given the extreme dearth of this type of data in the sector, where to start is most likely the issue we face. While the importance and relevance of measuring contributions to such things as social capital is undisputed, the first order priority might best be confined to some crude measures. The selection of these measures would benefit from a consultation with the sector.

The research design and methodological issues that remain unresolved for the larger voluntary health sector project have a major bearing on the selection of variables to be measured and, moreover, on how the data are brought to bear on the research objectives. For example, is the goal of the project to produce generalizable results, that is, do we need a statistical sample? Are there other methods and designs that would allow us to extrapolate and make statements concerning the size, scope and value of the voluntary health sector in Canada? The answers to these questions will definitely provide a clearer understanding of which measures would be best suited to certain research designs and methodologies.

A number of options and blends of options might be considered. While it is beyond the scope of this paper to deal with them in detail, we have opted to discuss some of them in order to demonstrate their relationship to the choice of measures and feasibility of collecting data.

The typology will inform on the boundaries of the sector and it could be assumed that the geographical universe is Canada. The

methodological issues associated with the sample frame most likely will require technical strategies in order to overcome issues such as the potential inability to derive a statistical sample. These strategies will, in turn, have a bearing on the measures or data collected.

For example, if a multi-pronged methodology were developed that focussed on the national and provincial level of the sector (that is, national and provincial organizations), and a sample of cities (small rural or isolated, mid-sized and large metropolitan), then it may be possible to construct a picture and measures of the geographical universe through extrapolation and other techniques. The municipal study approach could be patterned after or adapted to the component modelling approach used to estimate national incidence and prevalence of HIV in the United States (Holmberg, 1996). Within the context of the sector, the component model approach might involve identifying "natural laboratories" such as Peterborough or Sherbrooke where the compact size and factors concerning generalizability have allowed major corporations to conduct large marketing studies and surveys. This also assists in containing the size of the sample and promotes a more manageable dataset.

Case studies at the national, provincial or municipal level would provide the opportunity for gaining insight ("drill-down") into more robust measures of inputs, outputs and outcomes.

9.2 What Should We Measure

The research design, methodology and feasibility issues offer a backdrop for creating a menu or schedule of potential data elements that could be brought to bear on the larger question of assessing the size, scope and value of the voluntary health sector in Canada. Possible data elements for collection are presented below in Table 2 (a simplification of Table 1). It is important to recognize that this is a preliminary compilation and that a sectoral consultation and peer review process is still required to gain technical and sectoral validity. Within this process, discussion and recognition of the emergent approach to nonprofit effectiveness evaluation will be important. As well, the consultations must include a discussion of the ethical issues regarding information disclosure presented earlier. The feasibility of collecting input through output data is high, but low for collection of outcome data and contributions to the societal domain. To address the latter situation, we propose in Table 2 a simplification of the contributions to the societal domain of volunteers and nonprofits. For volunteers, a measurable impact is the personal account of the benefits of volunteering. Regarding broader societal impacts, the social accounting method applied by the Steering Committee (1998) to develop social return-on-investment ratios warrants more detailed examination. At a minimum, the method has the potential of supplying a common bottom-line indicator per organization and/or activity whose components may be additive across the voluntary health sector.

Table 2: Proposed Data for Determining the Value of the Voluntary Health Sector

	ORGANIZATIONAL ACTIVITY DOMAINS						SOCIETAL DOMAIN
	Fundraising	**Research**	**Service**		**Advocacy**	**Regulation**	
			Prevention/ health promotion/ education	**Care**			
Inputs – resources	- volunteer hours; paid staff hours; volunteer training hours; volunteer training expenses; program and service budgets						
Processes	- descriptive information about the types of programs and services offered						- volunteers' evaluations of personal benefits
Outputs	- # donors - $ raised	- # publications - # citations - # presentations - findings	- # materials published - # information campaigns	- # clients/ patients served	- # policy people contacted - # communiques	- # licenses issued - # licenses revoked - # professional practice complaints/ investigations	
Outcomes	- diversity of donor base - sustainability of donor base - expansion of donor base	- evidence of uptake and application of findings by practitioners and public - other evidence of knowledge transfer	- % target population reached - changes in knowledge, attitudes and behaviour - effect on disease incidence	- maintenance of current health state - reduced rate of decline - cure rate	- evidence of policy changes - issues acknowledged or given profile - measurable improvements in quality of life - services added or adjusted to meet needs	- evidence of public trust in a health profession	- social return-on-investment

ENDNOTES

BIBLIOGRAPHY

Abbott, Douglas A. et al., "The Influence of a Big Brothers Program on the Adjustment of Boys in Single-Parent Families," *The Journal of Psychology*, 1997;131(2):143-156.

Alberta Heritage Foundation for Medical Research , *A Snapshot of the Level of Indicator Development in Alberta Health Authorities: Toward a Common Set of Health Indicators for Alberta (Phase One)*, Alberta Heritage Foundation for Medical Research, 1997, Edmonton.

Ames, Seth , "How to Measure Value of a Development Program," *Fund Raising Management*, July 1996;12-13.

Andrews, Howard F., "Helping and Health: The Relationship between Volunteer Activity and Health-Related Outcomes," *Advances*, 1990; 7(1):25-34.

Babins-Wagner, Robbie, "Effectiveness of Abuser Treatment in a Community Based Family Violence Treatment Program," unpublished paper, Calgary Counselling Centre.

Baker, G. Ross et al., "Healthcare Performance Measurement in Canada: Who's Doing What?" *Hospital Quarterly*, Winter 1998/1999; 2 (2): 22-26.

Bart, Christopher K., "Mission Statement Rationales and Organizational Alignment in the Not-for-Profit Health Care Sector," *Health Care Management Review*, 1998; 23(4):54-69.

"Before You Embark on a Behavioural Health Outcome Measurement Project," cited on-line January 19, 1999 at www.workgroup.org.

Birch, Stephen, "As a Matter of Fact: Evidence-Based Decision-Making 'Unplugged'," Working Paper 97-2, Centre for Health Economics and Policy Analysis, McMaster University, March. 1997, Hamilton.

Bowen, Paddy, "Trends in Volunteerism: A Canadian Perspective," *The Journal of Volunteer Administration*,Spring 1997;4-6.

Canadian Council on Health Services Accreditation , *Performance Indicator Initiatives across Canada: An Overview*, Council on Health Services Accreditation, 1977, Ottawa.

Brett, E. A., "Voluntary Agencies as Development Organizations: Theorizing the Problem of Efficiency and Accountability," *Development and Change*, 1993; 24:269-303.

Brown, Eleanor, "Assessing the Value of Volunteer Activity," *Nonprofit and Voluntary Sector Quarterly*, March 1999; 28(1):3-17.

Brown, Ken W. and Mary L. Fischer, "Assessment Measures: Management's Yardstick," *Assessment and Evaluation in Higher Education*, 1994;19(3):163-174.

Browne, Paul Leduc and Pierrette Landry (1996), *The "Third Sector" and Employment: Final Report to the Department of Human Resources Development*, Canadian Centre for Policy Alternatives, 1996, Ottawa.

Buckingham, K., "A Note on HYE (Healthy Years Equivalent)," *Journal of Health Economics*, 1993; 11.

Buxton, Martin and Steve Hanney, "How Can Payback from Health Services Research be Assessed?," *Journal of Health Services Research Policy*, January 1996; 1(1):35-43.

Canadian Comprehensive Auditing Foundation , *Canadian Hospitals: Accountability and Information for Cost-Effectiveness, An Agenda for Action*, Canadian Comprehensive Auditing Foundation, 1987, Ottawa.

Canadian Council on Social Development, "Community-based Health and Social Services: Conference Report," 1985; November 24 to 27.

Canadian Centre for Philanthropy, "Developing a Typology of Canadian Voluntary Organizations: Issues and Considerations for Assessing the Value of Voluntary Organizations Working in Health," unpublished paper, 1999, Toronto.

——— , "The Effects of Health Care Reform on Philanthropy in Canada," Final Report, Association for Healthcare Philanthropy in Canada and the Canadian Centre for Philanthropy, June 1996.

Canadian Institute for Health Information, "Comparative Analysis of the HMRI Database – 1990/91," Canadian Institute for Health Information, Toronto.

Canadian International Development Agency, "Results-Based Management in CIDA - Policy Statement," cited on-line January 28, 1999 at www.acdi-cida.gc.ca.

Canadian Public Health Association, *Health Impacts of Social and Economic Conditions: Implications for Public Policy*, Canadian Public Health Association, 1997, Ottawa.

Canadian Hemophilia Society, *A Hands-On Guide to Planning and Evaluation: How to Plan and Evaluate Programs in Community-Based Organizations*, National AIDS Clearinghouse, 1993, Ottawa.

Centre for Health Promotion, and Toronto District Health Council, *Nine Steps to a Health Promoting Integrated Health System (IHS)*, Centre for Health Promotion, 1997, Toronto.

Chang, Cyril F. and Howard P. Tuckman,"The Goods Produced by Nonprofit Organizations," *Public Finance Quarterly*, January 1996; 24 (1): 25-43.

Chinman, Matthew J. and Abraham Wandersman , "The Benefits and Costs of Volunteering in Community Organizations: Review and Practical Implications," *Nonprofit and Voluntary Sector Quarterly*, March 1999; 28 (1): 46-64.

Clement, Jan P., Dean G. Smith and John R. C. Wheeler, "What Do We Want and What Do We Get from Not-for-Profit Hospitals," *Hospital and Health Services Administration*, Summer 1994; 39(2): 159-178.

Clements, Jason and Johanna Kompas (1998), "Estimating the Value of Volunteering," cited on-line March 29, 1999 at www.fraserinstitute.ca.

Day, Kathleen M. and Rose Anne Devlin, "Measuring the Nonprofit Sector," in Ronald Hirshhorn (ed.), *The Emerging Sector: In Search of a Framework*, Study No. CPRN|01, Canadian Policy Research Networks, 1997, Ottawa.

——— , *The Canadian Nonprofit Sector*, Working Paper No. CPRN|02, Canadian Policy Research Networks, 1997, Ottawa.

——— , "Models for Organizing the Delivery of Personal Health Services and Criteria for Evaluating Them," *Milbank Memorial Fund Quarterly*, 1972; Part 2.

Donabedian, Avedis,. "A Guide to Medical Care Administration, Volume II: Medical Care Appraisal," *American Public Health Association*, 1969, New York.

Dow, Warren , "The Voluntary Sector – Trends, Challenges and Opportunities for the New Millennium," *Volunteer Vancouver*, September 1997.

Drucker, Peter E., "What Business Can Learn from Nonprofits," *Harvard Business Review*, July-August;989:88-93.

Drummond, M.F., G.l. Stoddard and G.W. Torrance,. "Methods for the Evaluation of Health Care Programmes," *Oxford Medical Publications*, 1987, Oxford.

Echenberg, Havi, "Outcome-based Assessment," Draft Report to the Panel on Accountability and Governance in the Voluntary Sector, unpublished paper, 1998, Ottawa.

Ecumenical Coalition for Economic Justice , "Out of the Shadows," *Economic Justice Report*, October 1997; 8(3) .

Elkin, R. and Mark Molitor, *Management Indicators in Nonprofit Organizations: Guidelines to Selection and Implementation*, School of Social Work and Community Planning, University of Maryland, 1984, Baltimore.

Evaluation Practice, Spring-Summer 1996; 17(2).

Forbes, Daniel P., "Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997," *Nonprofit and Voluntary Sector Quarterly*, June 1998; 27(2):183-202.

Fox, Jonathan , "Policy Arena – The World Bank and Social Capital: Contesting the Concept in Practice," *Journal of International Development*, 1997; 9(7):963-971.

French, Michael T. and Kerry Anne McGeary , "Estimating the Economic Cost of Substance Abuse Treatment," *Health Economics*, 1997; 6:539-544.

Frohlich, Norman and Cam Mustard, *A Regional Comparison of Socioeconomic and Health Indices in a Canadian Province*, Working Paper No. 42, Canadian Institute for Advanced Research Program in Population Health, 1995, Toronto.

Furlong, W. et al., "Multiplicative Multi-Attribute Utility Function for the Health Utilities Index Mark 3 (HUI3) System: A Technical Report," *CHEPA Working Paper Series No. 98-11*, 1998, McMaster University.

Gafni, A, S. Birch and A. Mehrez, "Economics, health and health economics: HYEs versus QALYs," *Journal of Health Economics*, 1993;11.

Gassler, Robert Scott, "Nonprofit and Voluntary Sector Economics: A Critical Survey," *Nonprofit and Voluntary Sector Quarterly*, Summer 1990; 19(2):137-149.

Hall, Michael et al., *Caring Canadians, Involved Canadians: Highlights from the 1997 National Survey of Giving, Volunteering and Participating*, Minister of Industry, 1998, Ottawa.

Hancock, Lynne et al., "Community Action for Health Promotion: A Review of Methods and Outcomes 1990-1995," *American Journal of Preventive Medicine*, 1997; 13(4):229-239.

Harriss, John and Paolo de Renzio, "An Introductory Bibliographic Essay," in "Policy Arena – Missing Link or Analytically Missing? The Concept of Social Capital," *Journal of International Development*, 1997; 9(7):919-937.

Hawrylyshyn, Oli., *The Economic Nature and Value of Volunteer Activity in Canada*, Secretary of State, 1978.

Hay, David I., "Evaluating Measures of Well-Being," Draft Working Paper, October 1998, The Information Partnership, 1998, Victoria, BC.

Hay, Robert D., *Strategic Management in Non-Profit Organizations: An Administrator's Handbook*, Quorum Books, 1990, Connecticut.

Hayes, Michael V. and James R. Dunn, *Population Health in Canada: A Systematic Review*, CPRN Study No. H|01, Canadian Policy Research Networks, 1998, Ottawa.

Health Communication Unit , *Evaluating Health Promotion Programs*, University of Toronto, 1997, The Centre for Health Promotion, Toronto

Hedrick, Terry, "The Quantitative-Qualitative Debate: Possibilities for Integration," in Charles Reichardt and Sharon Rallis (eds), "The Qualitative-Quantitative Debate: New Perspectives," *New Directions for Program Evaluation*, Spring 1994; 61:45-53.

Herman, Robert D. and Richard Heimovics, "Researching non-profit organizational effectiveness: whose criteria? A rejoinder to Osborne and Tricker," *Voluntas* 1996; 6(1):93-99.

Herman, Robert D. and David O. Renz, "Multiple Constituencies and the Social Construction of Nonprofit Organization Effectiveness," *Nonprofit and Voluntary Sector Quarterly*, June 1997; 26(2):185-206.

Herman, Robert D., "Methodological Issues in Studying the Effectiveness of Nongovernmental and Nonprofit Organizations," *Nonprofit and Voluntary Sector Quarterly*, Fall 1990; 19(3):293-306.

Hirshhorn, Ronald, "The Raison d'Être of Nonprofits" in Ronald Hirshhorn (ed.), *The Emerging Sector: In Search of a Framework*, Study No. CPRN|01, Canadian Policy Research Networks, 1997, Ottawa.

Hodge-Williams, John, Nancy H. Doub and Roberta Busky, "Total Quality Management (TQM) in the Non-Profit Setting: The Woodbourne Experience," *Residential Treatment for Children and Youth* ,1995; 12(3):19-30.

Holland, Thomas P. (1988), "The Effectiveness of Non-Profit Organizations," *The Journal of*

Applied Social Sciences, Spring/Summer 1988; 12(2):202-221.

Jackson, Douglas K. and Thomas P. Holland, "Measuring the Effectiveness of Nonprofit Boards," *Nonprofit and Voluntary Sector Quarterly*, June 1998; 27(2):159-182.

Jensen, Jane, *Mapping Social Cohesion: The State of Canadian Research*, CPRN Study No. F|03, Canadian Policy Research Networks, 1998, Ottawa.

Jun, Minjoon, Robin T. Peterson and George A. Zsidisin, "The Identification and Measurement of Quality Dimensions in Health Care: Focus Group Interview Results," *Health Care Management Review, 1998;* 23(4):81-96.

Kawachi, Ichiro and Bruce Kennedy, "Health and social cohesion: why care about income inequality?," *British Medical Journal,* 1997; 314:1037-1040.

Kazandjian, Vahe A., Parnell Wood and Jane Lawthers , "Balancing Science and Practice in Indicator Development: The Maryland Hospital Association Quality Indicator (QI) Project", *International Journal for Quality in Health Care*, 1995; 7(1) (1995): 39-46.

Knapp, Martin, Vasiliki Koutsogeorgopoulou and Justin Davis Smith , "Volunteer Participation in Community Care," *Policy and Politics*, 1996; 24(2):171-192.

Leeder, Stephen, "Social Capital and Its Relevance to Health and Family Policy," cited on-line February 22, 1999 at www.pha.org.au.

Longo, Daniel R., "The Measurement of Community Benefit: Issues, Options and Questions for Further Research," *The Journal of Health Administration Education*, Summer 1994; 12(3):291-317.

Lorig, Kate et al., *Outcome Measures for Health Education and Other Health Care Interventions*, Sage Publications, 1996, San Francisco, California.

Magura, Stephen and Beth Silverman Moses, *Outcome Measures for Child Welfare Services: Theory and Applications*, Child Welfare League of America Inc., 1986, Washington.

Holmberg, Scott, "The Estimated Prevalence and Incidence of HIV in 96 Large US Metropolitan Cities," *American Journal of Public Health*, 1996: 86(5);642-654.

Makosky, Lyle, "National Voluntary Organizations Working in Health and Health Canada: Framing a New Relationship between National Voluntary Organizations Working in Health and Health Canada," InterQuest, May 27, 1998, Draft.

Maxwell, Judith , "Social Dimensions of Economic Growth. ,"*Eric J Hanson Memorial Lecture Series*, Volume III, University of Alberta, 1996.

McColl, Alastair et al., "Performance Indicators for Primary Care Groups: An Evidenced Based Approach," *British Medical Journal*, November 1998; 317(14) :1354-1360.

McEwan, Kimberley L. and Douglas A. Bigelow, "Using a Logic Model to Focus Health Services on Population Health Goals," *The Canadian Journal of Program Evaluation,* 1997; 12(1):167-174.

McNamara, Carter, "Basic Guide to Program Evaluation," cited on-line January 19, 1999 at www.mapnp.org/library.

McNelis, Rosemary Hattrup and William E. Bickel, "Building Formal Knowledge Bases: Understanding Evaluation Use in the Foundation Community," *Evaluation Practice*, 1996; 17(1):19-41.

Milne, Derek and Linda Gibson, "Quality Assurance in the Voluntary Sector," cited on-line January 20, 1999 at www.emerald-library.com.

Mobley, Catherine, "Toward a New Definition of Accountability: Using Applied Ethnography as a Tool for Change in the Voluntary Sector," *Journal of Contemporary Ethnography*, April 1997; 26 (1): 75-97.

Moore, Cynthia W. and Joseph P. Allen, "The Effects of Volunteering on the Young Volunteer," *The Journal of Primary Prevention*, 1996; 17(2): 231-258.

Mosser, Gordon, "Clinical Process Improvement: Engage First, Measure Later," *Quality Management in Health Care*, 1996; 4(4):11-20.

Mulgan, Geoff and Charles Landry, *The Other Invisible Hand: Remaking Charity for the 21st Century*, Demos, 1995, London.

Newcomer, Kathryn E., "Using Performance Measurement to Improve Programs," in Kathryn E. Newcomer (ed.), "Using Performance Measurement to Improve Public and Nonprofit Programs," *New Directions for Evaluation*, No. 75, Jennifer C. Greene and Gary T. Henry (editors-in-chief), Jossey-Bass Publishers, 1997, San Francisco.

Nuffield Institute for Health, "About Health Outcomes," cited on-line on January 19, 1999 at www.leeds.ac.uk/nuffield/infoservices.

Nutbeam, Don, "Health Outcomes and Health Promotion – Defining Success in Health Promotion," *Health Promotion Journal of Australia*, 1996; 6 (2): 58-60.

Oakley, Ann et al., "Supporting Vulnerable Families: An Evaluation of Newpin," *Health Visitor*, May 1995; 68(5):188-191.

Office of the Auditor General (1999), *Value for Money Audit Manual*, Office of the Auditor General, January 1999, Ottawa.

Osborne, Stephen and Mike Tricker, "Researching non-profit organizational effectiveness: A comment on Herman and Heimovics," *Voluntas*, 1996;6(1):85-92.

Panel on Accountability and Governance in the Voluntary Sector, "Building on Strength: Improving Governance and Accountability in Canada's Voluntary Sector," Final Report, February 8, 1999, Ottawa.

Patton, Michael, "A World Larger than Formative and Summative," *Evaluation Practice*, 1996; 17(2) (Spring-Summer):131-135.

Plantz, Margaret C., Martha Taylor Greenway and Michael Hendricks, "Outcome Measurement: Showing Results in the Nonprofit Sector," 1997, cited on-line January 5, 1999 at www.unitedway.org/outcomes.

Muttart Foundation, "Funding Programs: Agency and Program Evaluations," cited on-line on January 22, 1999 at www.muttart.org/agency.htm.

Plummer, John, *How Are Charities Accountable: A Study of the Approaches to Governance and Accountability Developed by Twelve Major Charities in Britain*, Demos, 1996, London.

Points of Light Foundation, American Cancer Society, National Health Council, *Will the Best of Intentions Be Enough?*, Points of Light Foundation, 1996, Washington.

Putnam, R , *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, 1993, Princeton, NJ.

——— , "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America," *Political Science & Politics*, 1995; 28(4).

Putzel, James, "Policy Arena – Accounting for the 'Dark Side' of Social Capital: Reading Robert Putnam on Democracy," *Journal of International Development* 1997; 9(7): 939-949.

Quality Standards Task Group, "A 'White Paper' on Quality in the Voluntary Sector," Quality Standards Task Group, 1998, London, England.

Reed, Paul, "Defining the Nonprofit Sector in Canada: What Questions Should We Be Asking," in Ronald Hirshhorn (ed.), *The Emerging Sector: In Search of a Framework*, Study No. CPRN|01, Canadian Policy Research Networks, 1997, Ottawa.

Rekart, Josephine, *Public Funds, Private Provision: The Role of the Voluntary Sector*, UBC Press, 1993, Vancouver.

Richmond, Betty Jane, "Counting on Each Other: A Social Audit Model to Assess the Impact of Nonprofit Organizations – A Summary," PhD. Thesis, Ontario Institute for Studies in Education, University of Toronto, 1999, Toronto.

Rifkin, Jeremy, *The End of Work: The Decline of the Global Labour Force and the Dawn of the Post-*

Market Era, G.P. Putnam's Sons, 1995, New York.

Rissel, Chris, Jeanette Ward and Peter Sainsbury, "An Outcomes Approach to Population Health at the Local Level in NSW: Practical Problems and Potential Solutions," *Australian Health Review*, 1996; 19(2):23-39.

Rose-Ackerman, Susan, "Altruism, Nonprofits and Economic Theory," *Journal of Economic Literature*, June 1996; 34:701-728.

Ross, David P., "The Economic Value of Volunteer Work," *Transition*, June 1994;16-17.

____ , *Economic Dimensions of Volunteer Work in Canada*, Department of Secretary of State, 1990, Ottawa.

Rossi, Peter H. and Howard E. Freeman (1993), *Evaluation: A Systematic Approach 5*, Sage Publications, 1993, Newbury Park, CA.

Roundtable on Comprehensive Community Initiatives for Children and Families, *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts*, James P. Connell, Anne C. Kubisch, Lisbeth B. Schorr, Carol H. Weiss (eds), The Aspen Institute, 1995, Washington.

____ , *New Approaches to Evaluating Community Initiatives: Volume 2, Theory, Measurement and Analysis*, Karen Fulbright-Anderson, Anne C. Kubisch, James P. Connell (eds), The Aspen Institute, 1998, Washington.

Schmid, Hillel, "Non-Profit and For-Profit Organizations in Home Care Services: A Comparative Analysis," *Home Health Care Services Quarterly*,1993; 14(1):93-112.

Scott, Jacquelyn Thayer, "Defining the Nonprofit Sector," in Ronald Hirshhorn (ed.), *The Emerging Sector: In Search of a Framework*, Study No. CPRN|01, Canadian Policy Research Networks, 1997, Ottawa.

Rodriguez-Spagnolo and Rosana Georgia, "A Study of Program Evaluation Capacity in Nonprofit Organizations (Michigan)," PhD Dissertation, University of Michigan, 1992.

Scriven, Michael (1996), "Types of Evaluation and Types of Evaluator," *Evaluation Practice*, Spring-Summer 1996; 17(2):151-162.

Secretary of State, *The Development of a Typology of the Voluntary Sector in Canada*, Policy Coordination, Analysis and Management Systems Branch, December, 1983.

——— , *People in Action*: Report of the National Advisory Council on Voluntary Action to the Government of Canada, 1977*a* September.

——— , *People in Action*: A Sampler and Summary of a Report by the National Advisory Council on Voluntary Action, 1977*b* October.

Sharpe, David , *A Portrait of Canada's Charities: The Size, Scope and Financing of Registered Charities*, Canadian Centre for Philanthropy, 1994, Toronto.

Sheehan, Robert M., Jr., "Mission Accomplishment as Philanthropic Organization Effectiveness: Key Findings from the Excellence in Philanthropy Project," *Nonprofit and Voluntary Quarterly*, March 1996; 25 (1) : 110-123.

Shomer, Robert, "Before you Embark on a Behavioural Health Outcome Measurement Project," cited on-line at www.workgroup.org.

Smith, David Horton, "Grassroots Associations are Important: Some Theory and a Review of the Impact Literature," *Nonprofit and Voluntary Sector Quarterly*, September 1997; 26(3): 269-306.

Smith, David Horton and Ce Shen, "Factors Characterizing the Most Effective Nonprofits Managed by Volunteers," *Nonprofit Management and Leadership*, Spring 1996; 6(3) (Spring):271-289.

Social Planning Council of Ottawa-Carleton, *Doing Less with Less: Report on the 1997 Ottawa-*

Carleton Community Agency Survey, Spring 1998.

Steering Committee, Social Accounting Framework Project, *Counting on Nonprofits: The Final Report of the Social Accounting Framework Project, April 1998*, United Way of London and Middlesex, 1998, London.

Sudman, Seymour and Norman M. Bradburn, *Asking Questions: A Practical Guide to Questionnaire Design*, Jossey-Bass Publishers, 1982, San Francisco, California.

Thompson, Joy C., "Program Evaluation within a Health Promotion Framework," *Canadian Journal of Public Health*, March-April 1992; Supplement 1:S67-S71.

Todd, Rebecca and Kavasseri V. Ramanathan, "Perceived Social Needs, Outcomes Measurement and Budgetary Responsiveness in a Not-for-Profit Setting: Some Empirical Evidence," *The Accounting Review*, January 1994; 69(1):122-137.

Torrance, GW, WH Thomas and DL Sackett, "A Utility Maximization Model for Evaluation of Health Care Programs," *Health Services Research*, Summer 1972; 7, no 2 .

Torrance, GW et al., "Multi-Attribute Preference Functions for a Comprehensive Health Status Classification System.," *CHEPA Working Paper Series No. 92-18*, McMaster University, 1992.

Torres, Cruz C., Mary Zey and Wm. Alex McIntosh August, "Effectiveness in Voluntary Organizations: An Empirical Assessment," *Sociological Focus*, 1991; 24(3):157-174.

United Way of America, *Measuring Program Outcomes: A Practical Approach*, 1996, Alexandria, Virginia.

——— , "Example Outcomes and Outcome Indicators for Various Programs," 1997, cited on-line on January 5, 1999 at www.unitedway.org/outcomes/examples.html.

____ , "Update: Outcome Measurement Activities of National Health and Human Service Organizations," April 1998, Alexandria, Virginia.

Taylor, Martha E. and Russy D. Sumariwalla, "Evaluating Nonprofit Effectiveness: Overcoming the Barriers," in Dennis Young et al. (eds), *Governing, Leading and Managing Nonprofit Organizations*, Jossey-Bass Publishers, 1993, San Francisco, California.

Thomas, John Clayton, "Program Evaluation and Program Development," in Robert D. Herman and Associates (ed.), *The Jossey-Bass Handbook of Nonprofit Leadership and Management*, Jossey-Bass Publishers, 1994, San Francisco, California.

Van Til, Jon, *Mapping the Third Sector: Voluntarism in a Changing Social Economy*, The Foundation Centre, 1988, New York, New York.

Veenstra, Gerry (1995), "Social Capital: What is it?," Centre for Health Economics and Policy Analysis, Department of Sociology, McMaster University, 1995, Hamilton.

Voluntary Sector Roundtable, "Backgrounder: Accountability in the Voluntary Sector," 1997, Ottawa.

Volunteer Canada, "The importance of voluntary activity in Canada," cited on-line January 21, 1999 at www.volunteer.ca.

____, "The potential of volunteer work for developing skills," cited on-line January 21, 1999 at www.volunteer.ca.

____, "Volunteering is a two-way street," cited on-line January 21, 1999 at www.volunteer.ca.

____ (1998), "Volunteerism in the Canadian Health Sector," Ottawa: Volunteer Canada, September.

Waitzer, Edward J., "Philanthropy and Prosperity – Rebuilding Social Capital," cited on-line January 29, 1999 at www.charityvillage.com.

Waysman, Mark and Riki Savaya, "Mixed Method Evaluation: A Case Study," *Evaluation Practice*, 1997; 18(3):227-237.

Webber, Henry S., Book Review of *Accounting – A Social Institution: A Unified Theory for the Measurement of the Profit and Nonprofit Sectors*, by Julius Cherny, Arlene R. Gordon and Richard J. L. Herson, Quorum Books, 1992, New

York, in *Social Service Review*, 1994; 68:285-286.

Wholey, James, "Formative and Summative Evaluation: Related Issues in Performance Measurement," *Evaluation Practice*, Spring-Summer 1996; 17(2):145-149.

Wilkins, R and Owen Adams, "Healthfulness of Life: A Unified View of Mortality, Institutionalization, and Non-Institutionalized Disability in Canada," *Institute for Research on Public Policy*, 1983, Montreal.

Wilkinson, Richard G., "Socioeconomic Determinants of health – Health inequalities:

World Bank, "Designing Project Monitoring and Evaluation," *Lessons & Practices*, 1996; (8) (June), cited on-line January 19, 1999 at www.worldbank.org.

World Health Organization, "Definition of Health," cited on-line February 16, 1999 at www.who.int/aboutwho/en/definition.html.

Yates, Jessica, "Performance Management in Human Services," cited on-line January 18, 1999 at www.welfareinfo.org.

Yeo, Michael, *Ethics and Economics in Health Care Resource Allocation*, Queen's University of Ottawa Economic Projects 93-07, Canadian Policy Research Networks Inc., 1993, Ottawa.

relative or absolute material standards?" *British Medical Journal*, 1997; 314: 591-595.

Williams, Arthur R. and Chris Kindle, "Effectiveness of Nongovernmental and Nonprofit Organizations: Some Methodological Caveats," *Nonprofit and Voluntary Sector Quarterly*, Winter 1992; 21(4):381-390.

Wolfson, Michael C., *Pohem – A Framework for Understanding and Modelling the Health of Human Populations*, Working Paper No. 11, : Canadian Institute for Advanced Research Program in Population Health, 1994.

Young, T. Kue, *Population Health: Concepts and Methods*, Oxford University Press, 1998, New York.

Young, Dennis R., "Accountability the key to keeping the nonprofit sector on course," cited on-line January 19, 1999 at www.charityvillage.com.

Zacharakis-Jutz, Jeff and Stanley Gajenayake, "Participatory Evaluation's Potential among Nonprofit Organizations," *Adult Learning*, July/August 1994.

Ziebell, Mary T. and Don T. DeCoster, *Management Control Systems in Nonprofit Organizations*, Harcourt Brace Jovanovich, 1991, San Diego, California.

ORGANIZATIONS CONTACTED

Berkeley Consulting Group, Toronto
Calgary Counselling Centre
Community Services Consulting, Edmonton
Health Communication Unit, The Centre for Health Promotion, University of Toronto
Heart and Stroke Foundation of Canada
Heart and Stroke Foundation of Ontario
Metro United Way, Halifax
Muttart Foundation, Edmonton
National YMCA, Toronto
United Way Community Services of Guelph-Wellington
United Way of Canada, Ottawa
United Way of Calgary
United Way of Edmonton
United Way of Victoria

---

1. Revenue Canada's registered charitable organizations include hospitals, health service organizations other than hospitals, health charitable corporations, health charitable trusts, and health organizations not classified in the above categories. Unregistered health nonprofits are organizations classified as health-related.

2. Regarding autonomy, nonprofits such as hospitals may receive substantial funding from government, but their operations are managed by independent boards of directors.

3. Knapp, Koutsogeorgopoulou and Smith (1996) define informal volunteering as unpaid work undertaken by an individual not through a group or organization, for another individual (excluding close relatives). Brown (1999) uses implicit reciprocity, personal responsibility and mediation by an organization as factors to distinguish between informal and formal volunteering.

4. Qualitative data can become quantitative through coding and frequency counts of qualitative information.

5. The NSGVP survey produced a robust data set with the potential for examining various aspects of volunteering in nonprofits to which value can be assigned.

6. All the approaches to outcome-based assessment share common tasks (The Panel on Accountability and Governance in the Voluntary Sector, 1999): they identify outcome goals; they identify or develop ways to measure progress toward or achievement of those goals usually using logic models that separate the service impact from other causative factors (McEwan and Bigelow, 1997); and they disseminate the assessments to stakeholders and use them for planning and monitoring.

7. The balanced scorecard approach has four key perspectives to performance indicator development that answer four key questions: the financial (how do we look to funders); the internal (how do clients see us); the client viewpoint (what must the organization excel at); and innovation and learning (can the organization continue to improve and

create value) (Canadian Council on Health Services Accreditation, 1997).

8. From a conversation with Michael Yeo, Canadian Medical Association.

9. Determining value beyond a bottom line (that is, some social benefit), referred to as social audit, has been undertaken by a small number of private sector for-profit corporations such as Ben and Jerry's and Body Shop (Echenberg, 1998).