

Change towards outcome based performance management

An Expedited Synthesis Final Report

April 2011

By

Professors Raisa Deber and Robert Schwartz,
with the assistance of Professors Ross Baker, Jan Barnsley, Andrea Baumann, Whitney Berta,
Brenda Gamble, Audrey Laporte, Fiona Miller, Tina Smith, and Walter Wodchis.

Students: Kathleen Gamble, Corrine Davies-Schinkel, Tim Walker

Project Manager: Kanecy Onate

Administrative Support: Christine Day

Decision-making Partner: Ontario Ministry of Health and Long-Term Care, Public Health
Practice Branch

Funded by the CIHR Expedited Knowledge Synthesis Program, ETP-100086

Principal Investigator: Raisa Deber, PhD
Department of Health Policy, Management and Evaluation
Faculty of Medicine, University of Toronto
Health Sciences Building
155 College Street Suite 425
Toronto, ON M5T 3M6
E-mail: raisa.deber@utoronto.ca

Change towards outcome based performance management
An Expedited Synthesis
Table of Contents

Executive Summary	i
Acknowledgments.....	iv
What were we asked to do?	1
Approach.....	1
Methods.....	1
Literature search.....	1
Conceptual framework.....	3
Definitions: Accountability, performance measurement, performance management and change management	3
Implementing performance measurement: Goals and indicators.....	5
The issue of measurement.....	6
Selecting indicators.....	7
Likely impacts.....	7
Tools for accountability	8
Analytical approach	10
What is public health	10
Defining the cases	13
The Performance Measurement Ladder.....	15
Findings.....	17
What works?	18
Appendix 1: Performance Measurement Ladder, Selected Public Health Activities	20
Appendix 2: Knowledge Transfer.....	24
Appendix 3: Selected case examples (see separate file).....	24
Appendix 4: References reviewed (see separate file).....	24
References Cited	25

Change towards outcome based performance management

By

Professors Raisa Deber and Robert Schwartz,

with the assistance of Professors Ross Baker, Jan Barnsley, Andrea Baumann, Whitney Berta, Brenda Gamble, Audrey Laporte, Fiona Miller, Tina Smith, and Walter Wodchis.

Students: Kathleen Gamble, Corrine Davies-Schinkel, Tim Walker

Project Manager: Kanecy Onate

Administrative Support: Christine Day

Decision-making Partner: Ontario Ministry of Health and Long-Term Care, Public Health Practice Branch

Executive Summary

This report synthesizes what we learned about the introduction and use of outcome based performance management systems for public health organizations. Our systematic literature review focused on characterizing outcome based performance management systems introduced and implemented in other jurisdictions for public health, and for other publicly financed health care programs. To the extent possible, we also examined what change management strategies were used to support shifts to outcome based performance measurement, as well as how well they worked, and under what contexts.

Approach and Methods

The search strategy used such keywords as: indicators, accreditation, Balanced Scorecard, evidence-based public health, local public health, performance measurement, performance standards, and public health management, alone and in combination. The database started with 213 references supplied by our decision-making partner. At present, there are 969 references entered for this project; they are listed in Appendix 4.

The unit of analysis is the case example, coded to allow comparison and extraction of best practices. The series of case examples are from: UK, Australia, New Zealand, US, and the Canadian provinces of Saskatchewan and BC. Examples are classified in terms of public health activities, using as the base the 2008 Ontario Public Health Standards Framework.

Performance measurement and management are often justified as a way to improve accountability. We concentrate on 4 currently-used approaches to accountability—Financial incentives, Regulations, Information directed towards patients/payers, and Professionalism/stewardship—which represent variations on the ‘expenditure,’ ‘regulation,’ and ‘exhortation’ governing instruments.

Each case example (described briefly in Appendix 3) captured the following information (as available):

Who was involved? (including involvement in paying for the service, providing the service, collecting the data, reporting the data, receiving and distributing the data, and

monitoring performance.) Depending upon the activity, there may be a number of actors involved, and coordination may be difficult.

What activities are being reported on? Production characteristics, in particular *contestability, measurability, and complexity*, affect how performance can be measured and managed. We have tried to clarify what they are reporting on, and to whom, as well as what they are not reporting on, and the goals of the exercise.

The nature of *implementation*, and how the information is being used for performance management: Who has access to the information? Who can use it?

Consequences - whether the jurisdiction appears to be using the measurement for performance management, and if so, what tools they are using, what happens when they succeed or fail, and whether there is any information about unintended consequences. This includes consideration of organizational effectiveness.

We developed a *performance management ladder* to compare cases on the characteristics of performance measurement and its use for performance management. Values on the ladder range from 0 to 4 (with 4 sub-divided into 4I and 4II.) Level 0 indicates that no goals were defined. At level 1, goals are defined but indicators have not been developed and data collection does not appear to be taking place. Level 2 denotes that goals and indicators are defined but data collection does not appear to be taking place. At level 3, goals and indicators are defined, and data are being collected, but there is no indication that the data are being used for management. (In this category, it is often unclear how widely disseminated the data is, whether data collection has been standardized, and whether a common performance measurement system is being used.) Level 4 indicates that goals and indicators are developed, data are being collected, and the data are being used for something (these data can be used differently through separate accountability tools). The subcategories of level 4 characterize how and where the data are being used.

Key Findings

1. Most of the examples relate to attempts to devise indicators that can be used by a variety of actors for a variety of purposes. In a small number of cases, these are linked to other policy instruments. In general, the case examples used performance measurement primarily as an information policy instrument. Some examples in which performance measurement is used to support expenditure-based approaches were found. There were also uses of performance measurement linked with regulation, particularly for traditional public health activities such as food and water safety or the safety of facilities caring for children.
2. We found very little information about the use of change management, both because many of the case examples focused on indicators/information rather than performance management, and because there were few published evaluations of the actual impact of those experiments.

3. Cases demonstrate a heavy emphasis on using indicators for surveillance. This often entails partnerships between government and outside non-governmental organizations (NGOs) for data collection and dissemination. In many of the cases, particularly in the UK, these arm's length agencies are used to compile indicator data from multiple sources, and make it readily available, often through web sites that facilitate comparison across organizations and/or jurisdictions. These results do not come without cost; a review of the "new public management" in health care, particularly in the UK, has noted an increase in administrative costs.
4. We found very few examples where the performance measurement data were being used for management. Particularly in the UK and New Zealand, there is an emphasis on using indicators for public reporting to allow different organizations to compare their performance. It is likely that some of these organizations in turn use that information for management purposes, but such internal activities rarely result in publications and hence many of these efforts would not appear in the literature search.
5. Public health clearly differs from the sort of clinical examples being used in other sub-fields in health care. One reason may be the disaggregated nature of public health delivery, and the difficulty of enforcing policy levers on partners. Other issues arise from the fact that much of traditional public health measures success in terms of events not experienced. This can be problematic for performance measurement for several reasons. First, when successful, public health is invisible – few are aware of disease outbreaks that did not occur. Second, there may be incentives to avoid looking for problems, or to misclassify them.
6. One striking finding is the extent to which use of performance measurement and management is focused on clinical services addressed to individuals. Activities directed towards improving the health of populations, particularly those with a preventive orientation, tend not to be included. Another trend, visible in the UK P4P, is for success to be measured as doing more of particular things (*e.g.*, screening tests, medication, some immunization) for particular populations (*e.g.*, people with chronic diseases); one unintended consequence is that processes for how services are provided receive more attention than outcomes. Another, particularly crucial for public health, is that prevention and population health may risk being lost in the shuffle.
7. One key positive impact of these efforts has been increased emphasis on data quality. Particularly where delivery is disaggregated, ensuring that all define their terms similarly, and collect and share data, is itself a major step forward.
8. We found differences between the development of performance measures as a management system, versus using measurement for improvement. When these measures are used as a management system, the literature suggests that the indicators need to be rigorously defined, and carefully measured to allow comparison across organizational units. On the other hand, where managers and teams are focused on improving performance, there is a need for local, "rough and ready" quality improvement measures.

9. Some key variables that appear to influence what is being included in performance measurement/management systems include:
 - Ease of measurement;
 - Data quality. Jurisdictions vary considerably in how good the data are. For example, Canada does not yet have good data about immunization at the national level;
 - Ability of organization to control outcomes (e.g., immunization vs. tobacco control);
 - Ability to measure success in terms of doing things (rather than preventing things);
 - What is already happening. One example is the UK P4P for physicians, which is generally considered to have been highly successful. However, there is some suggestion that what is being rewarded is better recording, rather than changes in practice. The indicator systems appear to, in part, reward providers for things they were already doing, which in turn raises questions about who gets to set the indicators.

One important caveat for any performance measurement/performance management system is that it does not, and cannot, capture all activities. In that connection, it is important to recognize that most providers are professionals, who want to do a good job. Performance measurement/management is only one component, but can give tools to allow all stakeholders to know how they are doing, and to improve performance.

To the extent that most providers wish to do a good job, the availability of good data to enable benchmarking and improvement is an important step forward. To the extent that the health of a population is dependent on multiple factors, many beyond the mandate of the health care system (both personal health, and public health), however, too extensive a reliance on performance measurement may risk unintended consequences of marginalizing critical activities. As ever, balance is key.

Acknowledgments

We very much appreciate the contributions of our partners at the Ontario Ministry of Health and Long Term Care, particularly Michèle Harding, Paulina Salamo, Sylvia Shedden, Monika Turner, and Michele Weidinger. The synthesis was funded by the CIHR Expedited Knowledge Synthesis Program, ETP-100086.

Change towards outcome based performance management

By

Professors Raisa Deber and Robert Schwartz,

with the assistance of Professors Ross Baker, Jan Barnsley, Andrea Baumann, Whitney Berta, Brenda Gamble, Audrey Laporte, Fiona Miller, Tina Smith, and Walter Wodchis.

Students: Kathleen Gamble, Corrine Davies-Schinkel, Tim Walker

Project Manager: Kanecy Onate

Administrative Support: Christine Day

Decision-making Partner: Ontario Ministry of Health and Long-Term Care, Public Health Practice Branch

What were we asked to do?

The project sought to identify the use of change management strategies, including quality improvement tools and strategies, that are needed to support the introduction of outcome based performance management systems. The project was asked to look at the use of strategies in jurisdictions comparable to Ontario in both public health sectors and larger health care delivery sectors (*e.g.*, regional health authorities, hospital reform exercises, etc.) to see what has been tried, what works, and where the thinking is going.

Approach

We took a staged and iterative approach, beginning with conversations with our decision making partner to clarify the scope of the study. The main approach was a systematic review of the literature on the defined topics. The review was designed to:

- identify and characterize outcome based performance management systems introduced and implemented in other jurisdictions for public health, and other publicly financed health care programs;
- learn what change management strategies were used to support (successful and less successful) shifts to outcome based performance measurement in these cases; and
- learn which change management strategies have worked well and less well and under what contexts.

Methods

Literature search

In discussions with the partners in defining the scope of the project, it was agreed that the question to be answered by this synthesis is not well suited to a Cochrane-style review; few jurisdictions would introduce outcome-based performance management systems using a randomized design. As the ESRC UK Centre for Evidence Based Policy and Practice has noted, social science reviews differ from the medical template in that they rely on a “more diverse pattern of knowledge production,” including books and grey literature.¹ We accordingly employed policy synthesis methodology,²⁻⁸ using a ‘best-evidence synthesis.’⁹⁻¹⁴ to allow the

review to recognize the difference between the methodology required for systematic reviews which address the question “what works?” (for which narrower methods may indeed be appropriate) and those which address a broader array of questions, such as “what combination of interventions works where, for which sub-populations, in which environmental circumstances, in which combinations, administered at what rate of intensity, over which period of time and in what order?”¹⁵

Our search strategy included multiple sources. To capture published and grey literature, we searched such databases as PubMed, Web of Science, and Google Scholar; these sites tend to capture different literatures, and thus helped ensure that key references were not missed.¹⁶ We also searched relevant web sites, both for the selected jurisdictions, and for the papers and reports produced by the World Health Organization (WHO), Organisation for Economic Co-operation and Development (OECD) and the European Observatory on Health Systems and Policies. We then analyzed both backwards and forward citation chains from key articles – that is, checking the relevant articles cited by that paper (backwards) and the materials citing that article (forward).¹⁷ Another helpful source was a US review of performance management in public health¹⁸ funded by the Robert Wood Johnson Foundation; they used the search terms United States[mh] AND (public health administration/st[majr] OR public health practice/st[majr:noexp] OR public health/st[majr:noexp]) AND (“performance standards” OR accreditation[majr:noexp] OR program evaluation[majr] OR outcome assessment(health care)[majr] OR process assessment(health care)[majr]). Their search yielded 90 articles and 44 grey literature documents related to public health, plus another 27 grey literature documents, as well as three peer-reviewed articles from outside public health relating to uses of performance management in government, primarily in the context of evaluation research. Note that there is a wide array of materials on their web site relating to performance management in public health (<http://www.turningpointprogram.org/toolkit/content/focus.htm#PerformanceManagement>), which was also reviewed for this synthesis.¹⁸⁻²² Another valuable source was the proceedings of a WHO European Ministerial Conference on Health Systems, which focused on performance measurement for health system improvement.²³

We followed the approach to systematic reviews recommended by Pawson et al.,²⁴ which, while still “explicit and transparent about the methods used,” takes a more iterative approach to developing the research questions, allowing the policy makers to participate in refining the research questions.²⁵ It recognizes that much of the analysis will, of necessity, be thematic and interpretative,^{3,26} including use of cross-case analysis^{6,8} and a ‘barriers and facilitators’ conceptual framework to draw together the conclusions.²⁷

The search strategy we employed used such keywords as: indicators, accreditation, Balanced Scorecard, evidence-based public health, local public health, performance measurement, performance standards, and public health management, alone and in combination. The abstracts were then scanned for relevance by our team. Using the Ontario Public Health Standards²⁸ as guidelines, the search strategy also sought to locate relevant case studies from Australia, New Zealand, the United Kingdom, the European Union, the United States, and Canada. Articles and other citable material were then entered into a computerized searchable bibliographic data base, using the free software package Papyrus. The database began with

material kindly supplied by Monika Turner (N=213), and was augmented. At present, there are 969 references entered for this project. Appendix 4 lists the references reviewed alphabetically.

The approach taken examined the general literature, and then selected key case examples to review. Case examples were selected by looking at the jurisdictions selected, with a focus on those that matched, corresponded, or contrasted with the Ontario Public Health Standards.²⁸ We also examined various reporting methods within each standard and then tried to determine what was being reported and why. We appreciate the input from our partners in confirming that the strategies and cases selected were appropriate to the goals of the synthesis. As noted below, as part of our analysis, we developed a Performance Measurement Ladder; the selected case examples have been placed on the ladder (Appendix 1) and are described briefly in Appendix 3.

Conceptual framework

Definitions: Accountability, performance measurement, performance management and change management

Increasing attention is being paid to the use of information to improve performance. Much of this dialogue is couched in terms of accountability.²³

Accountability is defined as having to be answerable to someone, for meeting defined objectives.²⁹⁻³¹ It has financial, performance, and political/democratic dimensions,³² and can be *ex ante* or *ex post*.³³ Within healthcare, this may translate into fiscal accountability to payers, clinical accountability for quality of care,³⁴ and/or accountability to the public. The actors involved may include various combinations of providers (public and private), patients, payers (including insurers and the legislative and executive branches of government), and regulators (governmental, professional); these actors are connected in various ways.^{35,36} The tools for establishing and enforcing accountability are similarly varied, and require clarifying what is meant by accountability, including specifying for what, by whom, to whom, and how. Performance management and measurement is frequently suggested as an important tool for improving systems of accountability.

As our review clarified, there is some variation within the literature and the cases examined in how various terms are defined, and in the purposes of the performance measurement exercise.³⁷ Underlying most of these examples is the sense that managing is difficult without measurement.³⁸

Performance measurement has been defined by the United States Government Accountability Office (GAO) as “the ongoing monitoring and reporting of program accomplishments, particularly progress toward pre-established goals.”³⁹ Their definition notes that such activities are typically conducted by the management of the program or agency responsible for them. The GAO contrasts this with program evaluation, which is often conducted by experts external to the program, and may be periodic or *ad hoc*, rather than ongoing. The GAO definitions, like many performance measurement systems in health care often use the framework of Donabedian, which focuses on various combinations of structures, processes, outputs, and outcomes.⁴⁰⁻⁴²

A number of approaches to performance measurement can be found in the literature.⁴³⁻⁴⁷ The focus of performance measurement systems can also vary, but increasing attention has been paid to using performance management as a way of improving system performance. Goals may also vary, but are often aligned with quality. A number of reviews of various efforts have been published; they include both examination of individual countries, and comparisons among OECD countries, including Canada, the US, the UK, and Australia.⁴⁸⁻⁵⁴ Much of the literature focuses on using performance measurement to improve clinical quality of care across a variety of settings, including primary care, and emergency care.⁵⁵⁻⁵⁷ Other projects focus on using performance measurement to improve governance, often using the language of accountability. For this to occur, ongoing data collection is important, so that management and stakeholders can use up to date information to monitor the quality of care being provided.⁵⁸ One approach is to use performance indicators.

Klazinga et al. link the development of national performance indicators for public health to the effort of the World Health Organization (WHO) to develop health targets. They suggest that “Terms such as safety, effectiveness, efficiency, equity, acceptability, accessibility, responsiveness, and adequacy are sometimes reformulated or modernized, but are still at the core of national frameworks for performance indicators that have been developed in the USA, Canada, UK, and Australia. Public health thinking is well represented in these frameworks on a (national) health system level but it is less obvious how this relates to indicators used for the monitoring and evaluation of health services.”⁵⁹

Performance management, in contrast, both paves the way for and requires a performance measurement system. Many measurement systems are developed with the goal of defining where improvements can be made, with the assumption that managers can employ them once the measurement results are examined.⁶⁰ Performance management can be defined as the action of using performance measurement data to effect change within an organization to achieve predetermined goals.⁶¹ *Change management* is defined by the US GAO as “activities involved in (1) defining and instilling new values, attitudes, norms, and behaviors within an organization that support new ways of doing work and overcome resistance to change; (2) building consensus among customers and stakeholders on specific changes designed to better meet their needs; and (3) planning, testing, and implementing all aspects of the transition from one organizational structure or business process to another.” (<http://www.gao.gov/special.pubs/bprag/bprgloss.htm>) There are a number of success stories in public management of using well-designed measurement systems to improve performance.⁶² As noted below, measurement may be necessary for management, but not all performance measurement systems assume that they will be used to manage.

This review focuses on the healthcare literature and experience, including but not exclusively public health. However, it is important to recognize that there is also a literature on the use of performance measurement and management in public administration, in both the public and private sectors.⁶³⁻⁶⁶ These authors place heavy emphasis on the role of organizational culture and political support in being able to implement change.

Implementing performance measurement: Goals and indicators

Chassin and colleagues⁶⁷ make several overall recommendations for creating a successful performance measurement system. They suggest that the number of areas being measured should be limited, the measurements need to be relevant to users of the system, system users should be involved in the development process, and data collection should be easily performed and within a time frame that allows for continual collection and data trending to occur. They also suggest that baseline data for measures should be collected before implementing any improvement strategies. Training should also be provided to users who are unfamiliar with data collection and performance measurement to ensure quality and consistency in collected data.

Accordingly, the first step to developing a successful performance measurement system is to clearly define what is going to be measured. McGlynn and Asch suggest that three considerations should be taken into account when choosing an area to measure: 1) how important the area of health care being measured is, 2) the amount of potential this area holds for quality improvement and 3) the degree to which health care professionals are able to control quality improvement in this area of health care. They define importance in terms of mortality/morbidity, but also utilization of health services and cost to treat.⁶⁸ Again, there is likely to be variation, depending on whether one is focused on particular patient groups, or on the health of the population. However, from the viewpoint of public health, these considerations point to the importance of surveillance systems, to provide decision makers with information about the prevalence of conditions, how they are being treated, and the outcomes of care.

Often implicit are what policy *goals* are being pursued. Different goals may imply different policies. Key goals of healthcare improvement are usually some combination of access, quality (including safety),⁶⁹ cost control/cost effectiveness, and customer satisfaction.^{70,71} Behn suggests objectives for accountability should be improved performance, fairness, and financial stewardship.⁷² This affects what organizations are accountable for. Often, policy goals may clash.⁷³ For example, hospital effectiveness may vary if measured in terms of doing more (increasing the number of admissions, market share, and/or occupancy rate), financial performance (net profit, cash flow), meeting the needs of the community (satisfaction of patients, providers), or delivering high quality care (better outcomes). An ongoing issue is the potential for unintended consequences if the measures selected do not reflect the full set of policy goals.⁷⁴ Indeed, one of the purposes of balanced scorecards is to make such potential conflicts between goals and measures more evident.⁷⁵⁻⁷⁸

Once an appropriate area of health care has been identified for measurement, the next step in developing a performance measurement system is to identify potential *indicators* that will be used in the measurement system. Indicators have been defined as “A measurement tool used to monitor and evaluate the quality of important governance, management, clinical, and support functions.”⁵⁹ Indicators define what is going to be measured and are quantified by a value that will indicate success by comparing that value to baseline data or benchmarked goals.⁷⁹ Indicators can be classified. For example, some authors assume that because performance must be measured against some specification, performance indicators do infer quality. Others (who do not necessarily represent a common view) distinguish between ‘Activity Indicators,’ which measure how frequently an event takes place, ‘Quality Indicators,’ which measure the quality of

care being provided, and ‘Performance Indicators,’ which do not infer quality but measure other aspects of the performance of the system (for example, the use of resources).⁸⁰ This in turn clarifies the importance of being clear about the goals one seeks to achieve (*e.g.*, various combinations of access, quality, and/or cost.).

Indicators are commonly stated as proportions; that is, they have a defined numerator and denominator. This allows comparison across jurisdictions. For example, an indicator examining childhood immunization would usually be defined as the total number of eligible children that received the particular vaccine, divided by the number of children eligible to receive it.

The issue of measurement

Loeb⁵⁸ argues that not everything in health care can or should be measured. In the clinical field, defining a number of relevant measurement indicators has led to successful performance measurement systems. Yet simply having one indicator in a measurement may not provide enough information. Alternatively, there is a risk of measures becoming unusable if there are too many indicators. Developing and running a performance measurement system is often expensive, and the data produced needs to be useful and interpretable for its users.

Many indicators are developed through a rigorous process. Typically, the first step involved in developing an indicator is a review of the literature.⁷⁹ Indicators are often developed from practice guidelines, practice evidence or decision rules.⁶⁸ In addition, the target population to which the indicator applies may need to be defined in terms of age (*e.g.*, upper and lower age limits to be included in data collection), population with diagnosed versus suspected conditions, and/or the demographics of patient care to include in the measure. Data sources also need to be identified when developing and choosing a set of indicators. The most common sources of data come from health care enrolment, administrative data, clinical data, and survey data.⁶⁸ Once a list of potential indicators has been generated from the literature review and target populations and data sources have been defined, the list of potential indicators will usually go through a review process to assess their reliability, validity, and cost-effectiveness.⁶⁸ Consensus panels are often used to perform item reduction and identify gaps in the list of indicators.⁷⁹ The next step is often to create a profile for each indicator that includes a definition of the indicator, the numerator and denominator measurement, specific data requirements and any risk-adjustment factors that will be taken into consideration.⁶⁸ By defining each indicator so specifically, the measurement system and its data collection processes can be implemented across different organizations/users in a consistent fashion. It helps to ensure that the data collected within the measurement system will be comparable and reliable across different users of the system.⁷⁹

Considerable efforts have been made to develop comparable indicators to enable cross-jurisdictional comparisons. These include the OECD quality indicators project,⁸¹ and the reporting standards for public health indicators.⁸² An offsetting concern is the recognition that strategic scorecards also must include locally relevant indicators. Achieving the right mix between local relevance and the ability to compare across organizations is crucial.

Selecting indicators

One ongoing issue is what sorts of indicators should be used. The factors important to those individuals providing clinical services to clients often differ from those important to program managers, payers, or health systems.⁸³ One class of indicators focuses on adverse outcomes, either at the individual level (*e.g.*, adverse events), or at the system level (*e.g.*, avoidable deaths). Klazinga et al. argue that “epidemiological research has shown the difficulties in validating [negative health outcomes] as indicators for the quality of care that was delivered.”⁵⁹

In selecting indicators, a key factor is the extent to which the elements affecting the measurement are under control of decision makers. Chassin and colleagues emphasize that for an outcome indicator to be relevant, it must be closely related to the health care processes that have an effect on the outcome.⁶⁷ In addition, there may be differences in what would be done with information; although the information may be valuable, it is difficult to hold managers accountable for things they cannot control. One obvious example is geography, which will often affect travel costs or access. Another, which affects population health, is the extent to which the various determinants of health (*e.g.*, income, housing, tobacco use, etc.) are under the control of public health organizations. Information may thus be helpful in affecting policy levers (*e.g.*, pricing of alcohol, tobacco) that other actors control, but less useful if program managers will be rewarded (or punished) for variables they cannot affect.

Other factors include whether different indicators are correlated (which can lead to double counting), how easy they are to measure (transaction costs), extent to which they are subject to “gaming” (*e.g.*, P4P in primary care – see case example A1) and whether they cover the outcomes of interest.^{84–89} Note that Ontario is one of many jurisdictions seeking to ensure more use of performance measurement and management. Indeed, this has been called for in reports by the Provincial Auditor.⁹⁰

Likely impacts

Another set of issues involves what will be done with the performance measures, including how they will be applied. Frequently, performance measurement involves setting performance targets, and assessing the extent to which these are being met. In turn, these may be used for funding (*e.g.*, results-based budgeting), and/or to identify areas for in-depth evaluation. External bodies may employ the information to ensure accountability. Managers may use them to monitor activities and make policies. Townley argues that “the use of performance measures reflects a belief in the efficacy of rational management systems in achieving improvements in performance.”⁷⁴ In the UK, use of fiscal levers is sometimes referred to as “targets and terror.”⁹¹

The way in which measures are likely to affect behaviour varies. Clearly, measurement is simplest if organizations produce a small number of services, have a limited number of goals, understand the relationship between inputs and results, and can control their own outcomes. As Townley notes, “A failure to ground performance measures in the everyday activity of the workforce is likely to see them dismissed for being irrelevant, unwieldy, arbitrary, or divisive.”

Other potential downsides are that “the time and resources taken to collect measures may outweigh the benefits of their use.”⁷⁴

A related set of factors relates to the organizational infrastructure.⁹² The workplace culture, including differences between the explicit goals, and what some have called the “implicit theories” or “theories in use” which affect day to day functioning, may affect the extent to which change initiatives are embraced, and performance changes.⁹³ This is in turn related to concepts of ‘street level bureaucracy,’ which deals with the extent to which it is simple to manage and observe the activities of those responsible for providing the given services.⁹⁴ Other less desirable organizational responses to performance measurement may include decoupling, a term used to refer to situations where specialist units are responsible for performance measurement, but where the measures have little impact on day-to-day activities and may lead to a sense that the measurement approach is ‘ritualistic’ and ‘bureaucratic’ rather than integral to improvement.⁷⁴ Even more alarmingly, measurement can lead to dysfunctional consequences, including focusing on measures rather than actual performance, impairment of innovation, gaming and creative accounting, potentially making performance worse.^{84,95} Other effects can be subtle; for example, placing less emphasis on prevention than on treating existing problems. The extent to which these positive or negative effects are realized may be heavily dependent upon context.

Tools for accountability

Performance measurement and management is often justified as a way to improve accountability. Our research team has recently begun a CIHR-funded Partnership for Health System Improvement (PHSI) project to ascertain the factors affecting the strengths and weaknesses of various approaches to accountability. The analytical framework we developed for that project identifies potential approaches to accountability and classifies these in terms of the political science concept often referred to as ‘policy instruments’ or ‘governing instruments.’⁹⁶⁻¹⁰¹ We concentrate on 4 currently-used approaches to accountability—Financial incentives, Regulations, Information directed towards patients/payers, and Professionalism/stewardship—which represent variations on the ‘expenditure,’ ‘regulation,’ and ‘exhortation’ governing instruments. In our PHSI research, we postulate that these approaches will have differing success when applied to various categories of services, and within various sub-sectors, with the likely outcomes depending upon: a) the policy goals being pursued; b) the governance/ownership structures and relationships in place, which in turn affect who will be accountable, and to whom; and c) the goods and services being delivered, and their production characteristics.

Our review suggests that the tools being used do reflect these approaches to accountability currently being employed in the health sector in Canada, and internationally:

1. *Financial incentives*, which adjust payments to induce providers to behave in the desired manner.¹⁰²⁻¹⁰⁵ These employ the ‘expenditure’ governing instrument. For example, Ontario’s hospital accountability agreements contain financial incentives for balanced budgets. The UK Pay For Performance (P4P) framework (case example A1) uses financial incentives.

2. *Regulations*, which employ the ‘regulation’ governing instrument,¹⁰⁶ play a major role in health care. These require providers to behave in a certain way. Although they can be backed

up through signing binding agreements, they may also rely on agency theory,^{107,108} and be enforced using professional regulatory bodies. Some are encompassed in provincial or federal legislation. Others are reflected in the accountability agreements being developed/used in Ontario for public health units, regional bodies (Local Health Integration Networks, or LHINs), and providers (e.g., hospitals, community service agencies), as well as to the regulations being used in the nursing home sub-sector. The literature notes the ongoing tension of balancing market forces and regulation.^{109,110} Related approaches may target the characteristics of the workforce, including credentialing,^{111,112} and/or the performance of the organization through accreditation.^{112,113} Our review confirmed that regulatory instruments are extensively employed in most jurisdictions for certain public health activities, such as clean water, and restaurant inspections (see, as some of many case examples, A6, A8, A9, F4).

3. *Information* directed towards potential users (patients, public and private payers) within a context of allowing market forces to work more effectively by encouraging rational choice of the ‘best’ care.^{114,115} These are one variant of the ‘exhortation’ governing instrument. This is a key component of the ongoing activities in *performance measurement and improvement*,^{51,116–119} including the PATH project, an international effort prepared by the GRIS group in Montreal for the European office of the WHO.¹²⁰ Issues include who establishes these measures, and who enforces them. Examples include report cards (e.g., for hospitals), audit reports, publicly-available inspection reports (e.g., for nursing homes) and quality indicators, including adverse events.⁶⁹ Note that this instrument may work indirectly (e.g., interest groups and media may affect the reputation of various providers, which in turn affect the willingness of patients and payers to purchase their services). This instrument may also not be relevant where natural monopolies exist (e.g., in rural/remote areas where the population is insufficient to support multiple competing hospitals). In the case of public health, this mechanism may be feasible for some services directed to individuals (e.g., immunization, health promotion), but not for traditional public health services provided by health units serving a distinct geographical area. The case examples in this report provide multiple examples of this instrument, often in the form of information posted on web sites (among other examples, see case examples A2, A3, A4, A7, A8, A9, B4, C1, C11, D2, D4, E3). However, in many cases, this information is directed towards providers rather than consumers, and would accordingly fall into the next category.

4. Reliance on ‘*professionalism*’ and ‘*stewardship*.’¹²¹ This approach employs a second variant of the exhortation governing instrument, but directs the information to providers rather than to payers or consumers.¹²² It relies on high trust and the expectation that providers – as a group – wish to do the right thing, but may need support in clarifying best practices. Clinical guidelines and other forms of evidence-based practice are currently being used in clinical practice. Continuing education can be used.¹²³ Under some circumstances, report cards may also fall into this category of approach, depending upon the indicators used and the dissemination approach adopted. Note that this approach is often backed up by regulatory approaches, including self-regulation of the professions. This instrument is widely used, including in public health.

An additional nuance is the extent to which blended models may be used. This is particularly evident in the use of additional policy instruments for enforcement, which may include various combinations of: Information (e.g., efforts to evaluate and improve the quality

of information, citizen engagement to widen the scope of inputs),¹²⁴ Expenditure (*e.g.*, fiscal incentives or penalties), Taxation (*e.g.*, tax breaks to encourage desired activities), and Regulation (*e.g.*, audit, accreditation, professional self-regulation, and legal sanctions).¹²⁵ Some jurisdictions have established formal appeal mechanisms for patients (*e.g.*, Norway has a Patient's Bill of Rights; Ontario has various review procedures for such issues as getting out-of-province coverage, a public complaints process for nursing homes, etc.). Other enforcement mechanisms may rely upon litigation (*e.g.*, malpractice, human rights). Auditor General reports fall within the Information category, but may often catalyze additional actions. A common thread in all of these accountability approaches is information; performance measures are critical enablers for financial incentives, regulation, and stewardship, as well as for affecting the types of choices made by patients and payers.

In general, the case examples we found used performance measurement primarily as an information policy instrument. Some examples in which performance measurement is used by expenditure-based approaches were found (particularly the UK Pay for Performance models in case A1). There were also uses of performance measurement linked with regulation, particularly for traditional public health activities such as food and water safety (A8, A9) or the safety of facilities caring for children (A4). However, many jurisdictions were focusing on assembling indicator information, with the expectation that those providing the services would wish to use them for performance improvement. Accordingly, a number of the case studies also focused on surveillance (especially communicable diseases and disaster preparedness) and emphasized the importance of public reporting and providing the public with updated information.

Analytical approach

The unit of analysis we employed is the case example, coded to allow comparison and extraction of best practices. The examples used are described briefly in Appendix 3, and categorized in Appendix 1. After consultation with our partners, we decided to classify these examples in terms of public health activities, using as the base the Ontario Public Health Standards Framework.²⁸ Note that there are some services found elsewhere which were not in the Ontario Framework; these are described in Appendix 3 but omitted from the table in Appendix 1 (*e.g.*, B11, Climate change). Note also that many of these activities not included in the Framework would still be performed in Ontario, just not through the organizations designated as public health. As the next section notes, public health encompasses a variety of activities, which are performed by a variety of actors in a variety of organizational settings.

What is public health

There are a number of different ways of describing what public health is, which often intermix services, processes, practices, and desired outcomes.¹²⁶⁻¹³³ Many, but not all, of its activities deal with the health of populations. Much, but not all, of the focus is on prevention. Some, but not all, of the activities are carried out by organizations designated as 'public health.' They include, but are not restricted to, the "health protection and promotion" functions, defined by Kirby and LeBreton as encompassing the following activities: "disease surveillance, disease and injury prevention, health protection, health emergency preparedness and response, health promotion, and relevant research undertakings."¹³⁴ For example, a 2005 BC report¹²⁹ noted a

series of diseases which are “currently held in check by public health programs.” Their examples included: water-borne disease (*e.g.*, cholera, typhoid), managed by water treatment and sewage disposal; food-borne diseases (*e.g.*, salmonella, hepatitis A, botulism), managed by regulating food preparation and sanitation; vaccine-preventable diseases (*e.g.*, smallpox, measles, diphtheria, polio, hepatitis, meningitis), managed by immunization programs; and new/emergent diseases (*e.g.*, HIV/AIDS, West Nile, BSE, SARS, multi-drug resistant tuberculosis), for which surveillance is crucial. However, the report went on to note that although the success of public health in controlling infectious diseases has led to a new focus on the chronic diseases contributing to the burden of illness and mortality, this has not always been reflected in public health research; instead, “research on public health services tends to focus on dated definitions of public health and is focused on infectious diseases and biological risk factors and behavioural change.”¹²⁹

Yet public health goes beyond even these functions. It can be seen as encompassing activities not usually falling within the mandate of public health agencies (*e.g.*, community economic development, education). It may also be heavily involved in provision of services to target populations (sometimes referred to as ‘indigent’ or ‘vulnerable’ persons), usually for populations or services not covered by health insurance systems (a role particularly important in the U.S.). In Canada, public health is often involved in providing some services to children and new parents, and preventive and clinical dental health services to vulnerable populations. In turn, some public health activities, including immunization, may be offered by public health, primary health care, and/or occupational health providers, depending on the jurisdiction and the service. A consultative process in Australia to classify public health activities noted that “The inclusion, or otherwise, of preventive services delivered on a one-to-one basis to individuals was particularly contentious. Such preventive services include screening and detection, immunisation, and counselling and lifestyle advice to support healthy behaviour, as well as management (through lifestyle changes or pharmacological means) of disease risk factors such as high blood pressure and high cholesterol.” There was particular disagreement as to whether non-communicable disease prevention should be classified as falling within public health, with the authors ultimately opting for an expansive definition. However, they noted that, in practice, boundaries in how services are organized can lead to difficulties in capturing information on services provided by non-health sectors (*e.g.*, transportation, welfare, education, environment).^{132,135}

It is important to recognize that a number of jurisdictions, including Ontario, have long sought to incorporate measures of performance. Indeed, Ontario’s Capacity Review process identified a need for accountability and performance management.¹³⁶⁻¹³⁸ One approach is to establish a set of ‘core programs’ for public health, which in turn can be linked to standards. For example, in Ontario, the Ontario Public Health Standards are published by the Minister of Health and Long-Term Care under the authority of section 7 of the Health Protection and Promotion Act.²⁸ There are also 26 program and topic specific Protocols that are incorporated into these standards.¹³⁹ An additional complexity is that responsibility for enforcing these standards is diffused across multiple ministries. The 2008 standards note that the Ministry of Health and Long-Term Care is responsible for administering the following standards: Foundational; Infectious Diseases; Environmental Health; and Emergency Preparedness, whereas the Ministry of Health Promotion had responsibility for: Chronic Diseases and Injuries and Family Health,

and the Ministry of Children and Youth Services for the administration of the Healthy Babies Healthy Children components of the Family Health standards.²⁸

Similar standards have been set by other provinces, including British Columbia (case example G1), Québec, and Saskatchewan (case example F5), while others are in development. As these processes have recognized, determining what public health programs are ‘core’ (and must be provided universally by all public health units in that province), which may be provided in some but not all jurisdictions, and which are best performed outside the public health system is complex. Ontario’s current process for identifying core programs speaks about four key principles: a) need (largely defined in terms of burden of disease); b) impact (largely defined in terms of having interventions for which reasonable evidence of effectiveness exists); c) capacity (including availability of resources and opportunity costs), and d) partnership and collaboration.²⁸ However, a 2006 review of public health funding, which included key informant interviews, indicated that these criteria were not necessarily used by local health units when determining what to offer. Precise decisions were dependent upon local circumstances, and included both decisions to go beyond the mandated requirements and offer programs valued locally (*e.g.*, dental services, animal services, services to frail seniors), as well as the potential that designating an activity as essential would not necessarily imply that it is carried out, particularly when resources are limited.¹⁴⁰ An ongoing issue is the tension between flexibility to allow programs to respond to local needs, vs. ensuring equity across the province/territory (or nation) in having access to defined levels of service. A related tension is whether core programs will be deemed minima, or maxima.

Given this wide array of activities falling under the public health rubric, it is not surprising that there is also considerable variability in who is involved in delivering public health, depending both on the service and the jurisdiction. There are often partnerships among levels of government, and with individual providers/provider organizations; the stakeholders may include: the public; employers, employees (including unions), and workplaces; individual providers, including physicians; hospitals, and their laboratories; regional authorities (which, in most provinces, have responsibility for hospitals); provincial/territorial governments (and, within them, departments with responsibility for health, health promotion, such other services as social services, environment, and labour, and public health labs); the federal government and other national agencies; and international bodies (*e.g.*, World Health Organization). Jurisdictions also vary in the extent of oversight over what their partners are doing. For example, Ontario’s influenza pandemic plan noted roles for: public health units; laboratories – including public health labs, teaching-hospital based labs, other hospital-based labs, and private labs in the community; community health services – defined as including the primary care providers, Community Care Access Centres, home care providers, long-term care homes, diagnostic centres, and other community-based agencies; emergency services; and acute care services.¹⁴¹

Although some jurisdictions, including Ontario, Québec and BC, have established provincial public health agencies, others have delegated public health to regional health authorities. A number of studies have argued that this can prove problematic, particularly when public health emergencies (*e.g.*, SARS) arise.^{142–149} These complexities mean that defining performance for public health will require careful attention to context, since there is likely to be

considerable variability as a function of the types of services being delivered. Realist synthesis stresses the importance of incorporating these types of characteristics.

Defining the cases

Selected case examples are included in Appendices 1 and 3. This section clarifies the information we attempted to extract when defining case examples, recognizing that there was considerable difference across cases in what was reported, and what was applicable. The case examples are briefly described in a mix of point and narrative form. Our reviews of these cases sought to establish the following information as available: Background; Who; What; Performance measurement; Implementation; Consequences; and Sources. Note that we found very little information about the use of change management, both because many of the case examples focused on indicators/information rather than performance management, and because there were few published evaluations of the actual impact of those experiments. The few examples which spoke of change management tended to be more oriented to strategy (how things should change over a given time period), rather than to how the data would be used to make such changes.

Background includes information to define what the case refers to, including: Jurisdiction(s) involved; Years (time period), recognizing that activities change over time; Organization(s) involved; and Services provided (linked, where possible, to the Ontario Public Health standards).

Who indicates the actors/stakeholders involved, with some attention to the *governance/ownership* structures in place,^{150–155} including the public-private mix in financing and delivery as it affects governance and accountability.^{156–160} Public refers to government at various levels— national, state/province, regional, municipal, as well as to quasi-government (e.g., most regional health authorities in Canada). Private includes not for profit; provider owned/small business; investor owned corporations, and individuals and families. For example, in Canada, much of public health is publicly financed and publicly delivered, whereas most hospital and physician services fall under what the OECD refers to as a “public-contract model,” in which public payers contract with private health-care providers.¹⁶¹ In turn, this means that many accountability arrangements are between government and the “third sector” (“civil society”), which presents additional complications.^{162–165} These in turn affect who will be accountable, to whom, and for what. As Denis et al. have noted, “Governance deals principally with the adaptation of organizations to new contingencies” and deals with “the roles of all regulatory, administrative, professional, and clinical authorities in the pursuit of collective goals.”¹⁵⁵ In that connection, dual accountability structures can add complexity.¹⁶⁶ It has been noted that the governance/ownership category of variables may be the most amenable to policy change.¹⁶⁷ Some models incorporate the concept of ‘soft governance’ approaches in which government “relies less on hierarchy than on information to steer local organizations.”¹⁶⁸ The literature suggests that governance/ownership can affect the ability to achieve and monitor such goals as quality improvement.^{169,170} A helpful resource is the framework developed by Denis *et. al* for the Canadian Council on Health Services Accreditation,¹⁵⁵ which identifies three governance models (agency, stakeholder, and stewardship).

Issues which policy makers would need to address under this category include, but are not restricted to:

- potential impact of ownership models. For example, the literature suggests considerable differences in the extent to which organizations would be expected to go beyond requirements as specified in contracts if they were not-for-profit (NFP), small business for-profit (FP/s) or investor-owned corporations (FP/c);^{73,159,171,172}
- legal responsibilities vs. practical enforcement; and
- enforcement mechanisms (criminal proceedings; Auditor General; Ombudsman; Privacy Commissioner; civil proceedings; appeal mechanisms, etc.).

Our case descriptions accordingly looked at who was involved (including the public/private mix) in the series of possible activities, including paying for the service, providing the service, collecting the data, reporting the data, receiving and distributing the data, and monitoring performance. Depending upon the activity, there may be a number of actors involved, and coordination may be difficult. Chronic disease management, for example, may be delivered by primary care, hospitals, public health, etc., and these groups may work independently. In such realms, one common approach we have found is for public or quasi-public organizations to develop indicators and provide information, in the hopes that other providers will use this to improve their performance, including placing more emphasis on preventive care.⁵⁹

The case examples also note *what* activities are being reported on in the performance measurement example, and (if available) how these relate to other activities being carried out by those organizations. Other elements in this category include how much autonomy the providers have in deciding what to do, and whether activities are mandatory or voluntary.

In that connection, our accountability framework notes the importance of examining the *production characteristics* of the goods and services, which will affect how performance can be measured and managed.^{159,173-175} In particular, we emphasize three, which the literature refers to as: *contestability*, *measurability*, and *complexity*.^{159,167,173-177} “*Contestable* goods are characterized by low barriers to entry and exit from the market, whereas non-contestable goods have high barriers such as sunk cost, monopoly market power, geographic advantages, and ‘asset specificity.’¹⁷⁶ *Measurability* relates to “the precision with which inputs, processes, outputs, and outcomes of a good or service can be measured.”¹⁷⁶ Monitoring performance is easiest when measurability is high. For example, it is relatively simple to specify the performance desired for conducting a laboratory test or collecting municipal garbage. In contrast, it would be more difficult to specify the activities to be expected of a general practitioner, and hence more difficult to monitor his/her performance and ensure quality. *Complexity* refers, not to how complex the particular goods and services are, but to whether the goods and services stand alone or require coordination with other providers. For example, laboratory tests are highly measurable, but gain much of their value by being embedded within a system of care, in which providers order tests appropriately and are aided in interpreting and acting upon their results. Additional insights arise from the theory of *transaction costs* and monitoring costs, which are also addressed in economic theories relating to incomplete contracting,^{178,179} which Williamson defines as “the comparative costs of planning, adapting and monitoring task completion under alternative governance structures.”¹⁸⁰

Another aspect noted is whether the particular activities apply to the full population, or to sub-sets of it. If so, how are the sub-sets defined? Examples include targeting certain income groups, age groups, diagnostic groups, etc.

We have tried to clarify what is being reported on, by whom, and to whom, as well as what they are not reporting on, and the goals of the exercise. In that connection, it is important to recognize that “management information is never value free.”⁵⁹ The choice of indicators reflects views about what is (and is not) important. It is also important to note who developed the indicators (and how), how the indicators are measured, how frequently data is collected, and how success is measured. The literature notes the importance of understanding unintended consequences, including how susceptible the indicators are to gaming. For example, much of traditional public health measures success in terms of events not experienced. This can be problematic for performance measurement for several reasons. First, when successful, public health is invisible – few are aware of disease outbreaks that did not occur. Second, there may be incentives to avoid looking for problems, or to misclassify them (*e.g.*, are injuries the result of accidents, or of family violence?). One perceived issue with P4P in the UK has been an incentive to over diagnose and overtreat.^{181,182}

We also note, to the extent this information is available from the published literature and web sites, the nature of *implementation*, and how the information is being used for performance management. Who has access to the information? Who can use it? We looked for information about problems, successes, costs, process of implementation, and any change management strategies used.

Finally, we noted *consequences* - whether the jurisdiction appeared to be using the measurement for performance management, and if so, what tools they were using, what happens when they succeed or fail, and whether there was any information about unintended consequences. This would include consideration of organizational effectiveness. Although it was beyond the scope of this review, one could see how this information might feed both into the financial management/audit functions, as well as into evaluation of how well boards of directors or management structures were functioning.

The Performance Measurement Ladder

Table 1 describes the “performance measurement ladder” we developed. In Appendix 1 and 3, we apply it to attempt to capture, for each of the categories listed in Ontario’s public health framework, the current approaches being used in the jurisdictions examined. Note that the categorization was often rough, and others might have classified certain programs differently. For example, we classified a variety of health promotion activities, including tobacco control, diabetes prevention, and obesity prevention, under Chronic Disease Prevention. We classified anti drug and anti alcohol abuse programs under Injury and Substance Misuse Prevention. Some dental health and hearing health programs, since they were directed towards children, were classified under Child Health. In rare events, multiple programs classified under the same framework category had different scores on the Ladder; in that case, we entered multiple scores. There were also a small number of programs that did not fit within the standards, and therefore are not on the chart, although they are described briefly in Appendix 3 (*e.g.*, Climate Control).

Fuller details about the precise programs are given in Appendix 3. Also note that we included some examples that had not yet progressed to program management, although they did help produce potential indicators, strategies, and/or data that could ultimately be used for those purposes. Note also that published case examples would generally have reached as least level 2 on the ladder. Rather than repeat long lists of indicators in Appendix 3, we summarize and indicate references, including links to documents where available.

We recognize that our findings are limited by what we found in our search strategy, since we could find only material published in publicly available writings. Note that the work plan did not include telephone interviews, although this could be accomplished in subsequent work. Note also that category 4 (the highest on the ladder) in turn has several sub-categories, some dealing with where the data is (4I) and some (not mutually exclusive) with how it is used (4II).

Table 1: The Performance Measurement Ladder

Level	Explanation
0	No goals defined for this category (as we can find through literature searches and websites)
1	Goals defined for this category, but no evidence that indicators have been developed; data collection does not appear to be taking place
2	Goals and indicators defined for this category, but data collection does not appear to be taking place
3	Goals and indicators defined, and data are being collected, but no indication that the data are being used for management. (In this category, it is often unclear how widely disseminated the data are, whether data collection has been standardized, and whether a common performance measurement system is being used.)
4	Goals, indicators are developed, data are being collected and the data are being used for something (these data can be used differently through separate accountability tools)
	Sub-categories of 4
4I	<i>Defines where the data are being used (hierarchy, where higher levels imply lower ones also hold)</i>
4I-a	The data stay within the independent organization
4I-b	The data are moving up through the levels of the organization or government
4I-c	The data are available publicly
4II	<i>Defines how the data are used within the organization that is collecting and using these data (the levels within Category 4II are not mutually exclusive)</i>
4II-d	The data are used in a management system – attached to money
4II-e	Regulatory approach (the data are used by regulators, <i>e.g.</i> , closing a restaurant or pool)
4II-f	Used within the independent organization to promote better practices or quality

Level	Explanation
4II-g	improvements Used as a management tool to compare and pull out poor or good performers through the public availability of the data, hoping organization would improve its performance when it compares itself against other jurisdictions

This report has drawn on a series of case examples from: UK, Australia, New Zealand, US, and several Canadian provinces, particularly Saskatchewan and BC. Appendix 1 places the case examples on the performance measurement ladder; Appendix 3 includes brief descriptions of each case.

Findings

One important caveat for any performance measurement/performance management system is that it does not, and cannot, capture all activities. In that connection, it is important to recognize that most providers are professionals, who want to do a good job. Performance measurement/management is only one component, but can give tools to allow all stakeholders to know how they are doing, and to improve performance. A second caveat is that we focused on published information; this may or may not reflect current activities in those jurisdictions. Successful interventions are also more likely to have been published.

Preliminary findings suggest considerable differences in what sorts of performance measurement and management are actually being done, not just by jurisdiction (which we expected), but by type of service. Looking at the performance measurement ladder, we found heavy emphasis on surveillance, and far less on explicitly using the indicator data for management. Additionally, there is more focus on processes of how services are provided than on outcomes.

A number of rationales are provided for this state of affairs. An excellent synthesis can be found in the edited volume with the proceedings of a recent WHO symposium.²³ It stresses the importance of clarifying causality, and the difficulty in holding providers accountable for outcomes that they cannot control. As one example they write that, “physicians working in socio-economically disadvantaged localities may be wrongly blamed for securing poor outcomes beyond the control of the health system.”²³ Risk adjustment methodologies can control for some, but not all, of this variation. Composite indicators can be useful, but only if transparent and valid. Similarly, it may be necessary to deal with random fluctuations before determining when intervention is needed to improve performance.

One striking finding that emerged from our review is the extent to which use of performance measurement and management is focused on clinical services addressed to individuals.²³ Activities directed towards improving the health of populations, particularly those with a preventive orientation, tend not to be included. As one example, the chapter in the report of the WHO symposium purportedly devoted to population health focuses almost exclusively on clinical treatment, including heavy focus on tracer conditions. One rationale given by these authors is that the performance measurement/management experiments they reported on wished to focus on the health care system. Their reaction to the fact that “it is often difficult to assess the

extent to which variations in health outcome can be attributed to the health system”¹⁸³ was accordingly to omit such measures. One concern arising from our review is that performance measurement approaches, by focusing so heavily upon the health care system *per se*, may skew attention away from important initiatives directed at improving the health of the population. Indeed, another chapter in the WHO symposium volume on “measuring clinical quality and appropriateness” explicitly states “A number of potential actions to improve population health do not operate through the health-care system (*e.g.*, ensuring adequate sanitation, safe food, clean environments) and some areas do not have health services that are effective in changing an outcome. Neither of these areas is fruitful for developing clinical process measures.”¹⁸⁴ Omitting such areas from measurement systems, however, may falsely imply that they do not matter.

Indeed, one of the papers reviewed took the example of the English public health strategy, *Health of the Nation*, launched in 1992, as a reason not to focus on the bigger picture.¹⁸⁵ They noted that the strategy’s impact on local policy making was seen to be negligible, in large part because the national authorities had few policy levers available with which to influence the performance (and priorities) of those local actors with responsibility for carrying out the relevant programs.

Our review stresses the importance of being aware of unintended consequences. For example, in the UK P4P, success tends to be measured as doing more of particular things (*e.g.*, screening tests, medication, some immunization) for particular populations (*e.g.*, people with chronic diseases); prevention and population health risk being lost in the shuffle.

Some key variables that appear to influence what is being included in performance measurement/management systems include:

- Ease of measurement
- Data quality. Jurisdictions vary considerably in how good the data are. For example, Canada does not yet have good data about immunization at the national level.
- Ability of organization to control outcomes (compare immunization to tobacco control)
- Ability to measure success in terms of doing things (rather than preventing things).
- What is already happening. One example is the UK P4P for physicians, which is generally considered to have been highly successful. However, there is some suggestion that what is being rewarded is better recording, rather than changes in practice. The indicator systems appear to, in part, reward providers for things they were already doing, which in turn raises questions about who gets to set the indicators.

What works?

One possible set of conclusions that can be drawn from this synthesis is guidance in identifying the ‘low hanging fruit.’ As shown in Appendix 1, there are some activities for which multiple jurisdictions have shown success, some where some have succeeded (and what distinguishes those), some where none have, and some where none appear to have tried to use performance measurement and management.

Most of the examples relate to attempts to devise useful indicators that can be used for a variety of purposes (*e.g.*, A2, A3, D3, E3).^{132,186} In a small number of cases, these are linked to

other policy instruments. For clinical services (see the case examples falling under Chronic Disease Prevention and Vaccine Preventable Diseases), particularly when these are delivered by GPs (as well as hospitals), there is some use of expenditure (see case example A1) and other financial incentives.¹⁸⁷ For food and water safety, there is considerable reliance on regulation (see case examples A6, A8, A9, F4). We could find very few examples of the less clinical services, however, where the performance measurement data was being used for management (few 4II-g examples); here, public health clearly differs from the sort of clinical examples being used in other sub-fields in health care. One reason may be the disaggregated nature of public health delivery, and the difficulty of enforcing policy levers on partners. One noticeable thrust appears to be using indicators for surveillance, which often entails partnerships between government and outside NGOs for data collection and dissemination. Particularly in the UK and New Zealand, there is also an emphasis on public reporting to allow different organizations to compare their performance. It is likely that some of these organizations in turn use that information for management purposes, but such internal activities rarely result in publications and hence many would not appear in the literature search.

One key positive impact of these efforts has been increased emphasis on data quality. Particularly where delivery is disaggregated, ensuring that all define their terms similarly, and collect and share data, is itself a major step forward. In some cases, this is linked to expansion of electronic health records.¹⁸⁸ Another characteristic found in many of the cases reviewed, particularly in the UK, is the use of arm's length agencies to compile indicator data from multiple sources, and make it readily available, often through web sites which facilitate comparison across organizations and/or jurisdictions. These results do not come without cost; a review of the "new public management" in health care, particularly in the UK, has noted an increase in administrative costs.¹⁸⁹

One theme is that there are differences between the development of performance measures as a management system, versus using measurement for improvement. When these measures are used as a management system, the literature suggests that the indicators need to be rigorously defined, and carefully measured to allow comparison across organizational units. On the other hand, where managers and teams are focused on improving performance, there is a need for local, "rough and ready" quality improvement measures.³⁷ The UK provides a number of examples, resulting from its heavy use of performance indicators. One review suggested that they had largely relied on external sources of control, linked with negative consequences for action, which in turn had given rise to conceptual and technical problems, particularly as related to the validity and reliability of the indicators, and the introduction of perverse incentives.¹⁹⁰

To the extent that most providers wish to do a good job, the availability of good data to enable benchmarking and improvement is an important step forward. To the extent that the health of a population is dependent on multiple factors, many beyond the mandate of the health care system (both personal health, and public health), however, too extensive a reliance on performance measurement may risk unintended consequences of marginalizing critical activities. As ever, balance is key.

Appendix 1: Performance Measurement Ladder, Selected Public Health Activities

A-1: International examples

	UK	Australia	NZ	Europe
Chronic Disease Prevention	4I-c,4II-d (A1) 4I-c,4II-g (A2) 4I-c,4II-f/g (A3)	3 (B1) 4I-c, 4II-g (B6) 3 (B8) 4I-c, 4II-f (B9) 3 (B10); 4I-c, 4II-f (B10) 3 (B11)	4I-c,4II-g (C1) 2 (C3) 2 (C4) 3 (C6) 3 (C7) 3 (C8) 3 (C9) 4I-c,4II-e (C12)	3 (D1) 3 (D2) 3 (D3) 4I-c,4II-f (D4)
Injury and Substance Misuse Prevention	4I-c,4II-f/g (A3) 4I-c,4II-e/g (A4)	3 (B1) 3 (B2) 2 (B11)	3 (C2) 3 (C8)	3 (D2) 3 (D3) 4I-c,4II-f (D4) 4I-c,4II-f/g (D9)
Reproductive Health	4I-c,4II-d (A1)			3 (D3)
Child Health	4I-c,4II-d (A1) 4I-c,4II-g (A2) 4I-c,4II-f/g (A3) 4I-c,4II-e/g (A4) 4I-c,4II-f/g (A5)	3 (B1)	3 (C7)	3 (D2)
Infectious Disease Prevention	4I-c,4II-d (A1) 4I-c,4II-g (A2) 4I-c,4II-e/f (A6)	3 (B1) 4I-c 4II-e (B4) 4I-c 4II-f (B9) 4I-c, 4II-f (B10) 4I-c, 4II-f (B11)	4I-c,4II-e/g (C10) 4I-c (C11)	3 (D2) 3 (D3) 4I-c,4II-f/g (D5)
Rabies Prevention	4I-c,4II-e/f (A6)		4I-c,4II e/g (C10)	
Sexual Health, STIs, and Blood-borne Infection Prevention	4I-c,4II-e/f (A6) 4I-c,4II-f (A7)	4I-c,4II-e (B4) 2 (B7) 4I-c 4II-f (B9) 4II-g (B10) 4I-c, 4II-f (B11)	4I-c,4II-e/g (C10)	4I-c,4II-f/g (D5) 4I-c,4II-f/g (D9)
Tuberculosis Prevention	4I-c,4II-g (A2) 4I-c,4II-f/g (A3) 4I-c,4II-e/f (A6)	4I-c,4II-e (B4)	4I-c,4II-e/g (C10)	3 (D2) 4I-c,4II-f/g (D5)
Vaccine Preventable Diseases	4I-c, 4-II-d (A1) 4I-c,4II-e/f (A6)	4I-c, 4II-f (B9) 4I-b, 4II-f (B10) 4I-b, 4II-f (B11)	4I-c,4II-g (C1) 4I-c,4II-e/g (C5) 4I-c,4II-e/g (C10)	4I-c,4II-f/g (D5)
Food Safety	4I-c,4II-e (A8)	4I-c 4II-f (B9)		

	UK	Australia	NZ	Europe
Water Safety	4I-b,4II-e/f (A9)	4I-c, 4II-e (B9) 4I-c, 4II-e (B10) 4I-c, 4II-e (B11)	4I-c,4II-e (C12)	4I-c,4II-g (D7)
Health Hazard Prevention	4I-c,4II-e/f (A6)	2 (B5) 2, 3 (B9) 2 (B10)	4I-c (C11) 4I-c,4II-e (C12)	3 (D3) 1 (D8)
Public Health Emergency Preparedness	4I-c,4II-e/f (A6)	3 (B1) 4I-c, 4II-e/f (B3) 3 (B10) 2 (B11)		

Classified by the Ontario Public Health Standards²⁸ and Performance Measurement Ladder. Cases described in Appendix 3.

UK Examples

- A1: Quality and Outcomes Framework – Pay for Performance (P4P system)
- A2: Compendium of Clinical and Health Indicators
- A3: Association of Public Health Observatories Health Profiles
- A4: Office for Standards in Education, Children’s Services and Skills (Ofsted)
- A5: Child and Maternal Health Observatory (ChiMat)
- A6: Health Protection Agency (HPA)- Disease Surveillance
- A7: Sexual Health Balanced Scorecard
- A8: Food Standards Agency
- A9: Drinking Water Inspectorate (DWI)

Australia Examples

- B1: National Public Health Partnership
- B2: National Injury Prevention and Safety Promotion Plan
- B3: Health Emergency Preparedness and Response
- B4: The Communicable Diseases Network Australia (CDNA)
- B5: National Environmental Health Strategy
- B6: National Tobacco Strategy
- B7: Sexual Health & Family Planning Australia Strategic Plan
- B8: Australian Chronic Disease Prevention Alliance
- B9: Western Australia Public Health Division
- B10: New South Wales Public Health Division
- B11: Victoria Department of Health

New Zealand Examples

- C1: Health Targets
- C2: Family Violence: Violence Intervention Program (VIP)
- C3: Quality Improvement Plan (QIP): Cardiovascular Disease and Diabetes
- C4: Diabetes
- C5: National Immunisation Register (NIR) and the Immunisation Advisory Centre (IMAC)
- C6: Nutrition (Healthy Weights)
- C7: Promoting Oral Health

- C8: National Drug Policy New Zealand
- C9: Cancer Control
- C10: Communicable Diseases
- C11: New Zealand Public Health Surveillance Report
- C12: Environmental Health Indicators Project (EHI)

European Union Examples

- D1: Cancer Indicators
- D2: ISARE
- D3: European Community Health Indicators and Monitoring (ECHIM)
- D4: EUROSTAT
- D5: ECDC (European Centre for Disease Prevention and Control)
- D6: Developing a national performance indicator framework for the Dutch Health System
- D7: WISE (Water Information System for Europe)
- D8: WHO Europe guidelines for indoor air quality: dampness and mould
- D9: EMCDDA (European Monitoring Centre for Drugs and Addiction)

A-2: North American examples

	USA	Saskatchewan	BC
Chronic Disease Prevention	4I-c,4II-g (E1) 4I-c, 4II-g (E3) 4I-c,4II-e (E4.1) 4I-c,4II-e (E4.2)	2 (F1) 4I-b,4II-f (F5) 2 (F6) 4I-c (F7)	3 (G1)
Injury and Substance Misuse Prevention	4I-c,4II-g (E1) 4I-c, 4II-g (E3)	2 (F1) 4I-b,4II-f (F5) 2 (F6)	3 (G1)
Reproductive Health	4I-c,4II-g (E1) 4I-c, 4II-g (E3) 4I-c,4II-g (E4.3)	2 (F1) 4I-b,4II-f (F5) 2 (F6)	3 (G1)
Child Health	4I-c,4II-g (E1) 4I-c, 4II-g (E3)	2 (F1) 4I-b,4II-f (F5) 2 (F6)	3 (G1)
Infectious Disease Prevention	4I-c,4II-g (E1) 4I-c, 4II-g (E3) 4I-c (E6)	2 (F1) 4I-b,4II-f (F5) 2 (F6) 4I-c (F7)	3 (G1)
Rabies Prevention		2 (F1)	3 (G1)
Sexual health, STIs, and Blood-Borne Infection Prevention	4I-c,4II-g (E1) 4I-c, 4II-g (E3)	2 (F1) 4I-b,4II-f (F5) 2 (F6)	3 (G1)
Tuberculosis Prevention	4I-c (E6)	2 (F1)	3 (G1)

	USA	Saskatchewan	BC
Vaccine Preventable Diseases	4I-c, 4II-g (E3) 4I-c (E6)	2 (F1) 4I-b,4II-f (F5) 2 (F6) 4I-c (F7)	3 (G1)
Food Safety	4I-c, 4II-g (E3) 4I-c,4II-e (E5)	2 (F1) 4I-c,4II-f (F3)	3 (G1) 4I-c, 4II-e
Water Safety		2 (F1) 2 (F2) 4I-c,4II-e/f (F4)	3 (G1)
Health Hazard Prevention	4I-c,4II-g (E1) 4I-c, 4II-g (E3) 4I-c,4II-e (E5)	2 (F1) 2 (F2)	3 (G1)
Public Health Emergency Preparedness		2 (F1)	3 (G1)

Classified by the Ontario Public Health Standards²⁸ and Performance Measurement Ladder. Cases described in Appendix 3.

US Examples

E1: Public Health Improvement Resource Center

E2: Performance Management Collaborative

E3: Healthy People 2010

E4: Chronic Disease Prevention and Health Promotion

E4.1: Behavioral Risk Factor Surveillance System

E4.2 CDC's Chronic Disease Indicators

E4.3 National Assisted Reproductive Technology Surveillance System (NASS)

E5: The Food Safety Office (FSO).

E6: Immunization Safety Office.

Saskatchewan

F1: Saskatchewan Public Health.

F2: Saskatchewan Environment

F3: Saskatchewan Health Food Safety Regulations

F4: Saskatchewan Safe Drinking Water Strategy

F5: Performance Management Accountability Indicators.

F6: Health Service and Outcome Indicators Project

F7: Comparable Health Indicators

British Columbia

G1: Core Public Health Functions Framework

Appendix 2: Knowledge Transfer

Representatives of our decision-making partner, the Public Health Practice Branch of the MOHLTC, have been briefed throughout the process of developing the knowledge synthesis plan, collecting information and developing the report. They joined the research team twice in discussions held during 2010 at the University of Toronto. Their input was extremely valuable in ensuring that the report was responsive to their concerns. Note that there were several changes in who the designated partner was. Initially, we worked with Monika Turner, who also provided some preliminary references. We then worked with Paulina Salamo, manager of the Practice and Standards Unit in the Public Health Practice Branch, Public Health Division, Michèle Harding, Manager (Acting) of the Performance Management Unit, Michele Weidinger in the branch and subsequently briefed the current Director of the Branch, Sylvia Shedden. This report was also reviewed and approved by members of the Branch before the final version was submitted.

This project has also been discussed in the course of meetings about the linked research project on Approaches to Accountability in Public Health. During the official launch of this study at University of Toronto in November 2010, representatives of MOHLTC, Peel Public Health and the Ontario Agency for Health Protection and Promotion held a meeting that discussed performance measurement and management as an essential component of accountability. Additional discussions took place at the Ontario Agency for Health Protection and Promotion and in a teleconference and face-to-face meeting with representatives of the Public Health Division. Emerging results from this knowledge synthesis were also shared in the course of informational interviews with seven senior key informants from two Ministries, OAHPP and Public Health Units. We will also supply our partner with a searchable electronic file listing all references entered for this project, including abstracts as available, plus pdf copies where copyright permits.

On Friday, November 19th, Deber presented findings from this study at the monthly rounds of the Strategic Training Program in Public Health Policy held at University of Toronto. The audience included policymakers, practitioners, faculty and students. Deber and Schwartz were invited to provide advice to an Ontario Ministry of Health and Long-Term Care led working group on allocation of resources to public health units, during which they presented emerging findings.

Abstracts on this study's findings have been accepted for presentation at the upcoming 2011 conferences of the Canadian Association for Health Services and Policy Research (CAHSPR) in Halifax, Nova Scotia and of the Canadian Public Health Association in Montreal, Québec.

Appendix 3: Selected case examples (see separate file)

Appendix 4: References reviewed (see separate file)

References Cited

- 1 Grayson L, Gomersall A. *A difficult business: Finding the evidence for social science reviews*. ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 19, University of London. 2003. <http://www.evidencenetwork.org/Documents/wp19.pdf>.
- 2 Lomas J. Using research to inform healthcare managers' and policy makers' questions: From summative to interpretive synthesis. *Healthc Policy* 2005 Sep;1(1):55–71.
- 3 Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesizing qualitative and quantitative evidence: A review of possible methods. *J Health Serv Res Policy* 2005 Jan;10(1):45–53.
- 4 Lavis J, Davies H, Oxman A, Denis J-L, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* 2005 Jul;10(1):35–48.
- 5 Lavis JN. Moving forward on both systematic reviews and deliberative processes. *Healthc Policy* 2006;1(2):1–5.
- 6 Mays N, Pope C, Popay J. Systematically reviewing quantitative and qualitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy* 2005 Jul;10(1):6–20.
- 7 Lavis JN, Davies HTO, Gruen RL, Walshe K, Farquhar CM. Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. *Healthc Policy* 2006;1(2):21–33.
- 8 Pope C, Mays N, Popay J. Informing policy making and management in healthcare: The place for synthesis. *Healthc Policy* 2006;1(2):43–48.
- 9 Slavin RE. Best evidence synthesis: An intelligent alternative to meta-analysis. *J Clin Epidemiol* 1995 Jan;48(1):9–18.
- 10 Whitemore R, Knafk K. The integrative review: Updated methodology. *J Adv Nurs* 2005;52(5):546–553.
- 11 Thorne S. The role of qualitative research within an evidence-based context: Can metasynthesis be the answer? *Int J Nurs Stud* 2009 Apr;46(4):569–575.
- 12 Hammersley M. Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence & Policy* 2005;1(1):85–100.
- 13 Forbes A, Griffiths P. Methodological strategies for the identification and synthesis of 'evidence' to support decision-making in relation to complex healthcare systems and practices. *Nursing Inquiry* 2002;9(3):141–155.

- 14 Popay J. *Moving beyond effectiveness in evidence synthesis: Methodological issues in the synthesis of diverse sources of evidence*. London UK: National Institute for Health and Clinical Excellence (NICE); 2006.
- 15 Roger R. A decision-maker's perspective on Lavis and Lomas. *Health Policy* 2006;1(2):49–54.
- 16 Walters WH. Google Scholar coverage of a multidisciplinary field. *Information Processing and Management* 2007 Jul;43(4):1121–1132.
- 17 Kuper H, Nicholson A, Hemingway H. Searching for observational studies: What does citation tracking add to PubMed? A case study in depression and coronary heart disease. *BMC Medical Research Methodology* 2006;6(4).
- 18 Public Health Foundation. *Performance management in public health: A literature review*. Prepared for the Performance Management National Excellence Collaborative. 2009. <http://www.turningpointprogram.org/toolkit/content/pmlitreview.htm>.
- 19 Lichiello P, Turnock BJ. *Guidebook for Performance Measurement*. Washington, DC. 2003. www.turningpointprogram.com.
- 20 Public Health Foundation. *Survey on performance management practices in states: Results of a baseline assessment of state health agencies*. Washington, DC. 2002. http://turningpointprogram.org/pages/pmc_state_survey.pdf.
- 21 Baker S, Ross H, Bialek R, Ali Z. *From silos to systems: Using performance management to improve the public's health*. Prepared by Public Health Foundation for the Performance Management National Excellence Collaborative. 2003. <http://www.turningpointprogram.org/toolkit/content/silostosystems.htm>.
- 22 Public Health Foundation. *Performance Management in Action: Tools and Resources*. Fifth in a series of Turning Point resources on Performance Management. 2004. <http://www.phf.org/infrastructure/PMtools>.
- 23 Smith PC, Mossialos E, Papanicolas I, Leatherman S. 1st ed. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, UK: Cambridge; 2009.
- 24 Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review- a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005 Jul;10(1):21–34.
- 25 Gough D, Elbourne D. Systematic research synthesis to inform policy, practice and democratic debate. *Social Policy and Society* 2002 Jul;1(3):225–236.
- 26 Pawson R. Evidence-based policy: The promise of 'realist synthesis'. *Evaluation* 2002 Jul;8(3):340–358.

- 27 Oliver S, Harden A, Rees R, Shepherd J, Brunton G, Garcia J, et al. An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation* 2005 Oct;11(4):428–446.
- 28 Minister of Health and Long-Term Care. *Ontario public health standards 2008*. Ottawa, Ontario. 2008.
www.health.gov.on.ca/english/providers/program/pubhealth/oph_standards/ophs/progstds/pdfs/ophs_2008.pdf.
- 29 Marmor TR, Morone JA. Representing consumer interests: Imbalanced markets, health planning and the HSAs. *Milbank Meml Fund Q Health Soc* 1980 Mar;58(1):125–165.
- 30 Emanuel EJ, Emanuel LL. What is accountability in health care? *Ann Intern Med* 1996 Jan 15;124(2):229–239.
- 31 Fooks C, Maslove L. *Rhetoric, fallacy or dream? Examining the accountability of Canadian health care to citizens*. Health Care Accountability Papers, No. 1: Health Network, Ottawa, ON: Canadian Policy Research Networks, March. 2004.
<http://www.cprn.com/documents/27403en.pdf>.
- 32 Brinkerhoff D. Accountability and health systems: Toward conceptual clarity and policy relevance. *Health Policy Plan* 2004;19(6):371–379.
- 33 Klein R. *An anatomy of accountability*. Overview Presentation for CHEPA's 6th Annual Health Policy Conference: 'The Buck Stops Where? Accountabilities in Health and Health Care', Hamilton, ON: McMaster University, 20 May. 1993.
- 34 Dobrow MJ, Sullivan T, Sawka C. Shifting clinical accountability and the pursuit of quality: Aligning clinical and administrative approaches. *Healthc Manage Forum* 2008 Fall;21(3):6–12.
- 35 Zimmerman SV. *Mapping legislative accountabilities*. Health Care Accountability Papers - No./5, Health Network, Ottawa, ON: Canadian Policy Research Networks, February. 2005. <http://www.cprn.org/documents/35118en.pdf>.
- 36 Shortt SED, Macdonald JK. Toward an accountability framework for Canadian healthcare. *Healthc Manage Forum* 2002 Summer;15(4):24–32.
- 37 Solberg LI, Mosser G, McDonald S. The three faces of performance measurement: Improvement, accountability, and research. *Int J Qual Health Care* 1997 Mar;23(3):135–147.
- 38 Gibberd R. Performance measurement: Is it now more scientific? *Int J Qual Health Care* 2005;17(3):185–186.
- 39 United States Government Accountability Office. *Performance measurement and evaluation: Definitions and relationships*. Washington, DC, April. 2005.

- 40 Donabedian A. Evaluating the quality of medical care. *Milbank Meml Fund Q Health Soc* 1966 Sep;44(3, Part 2):166–203.
- 41 Donabedian A. *The definition of quality and approaches to assessment*. Ann Arbor, Michigan: Health Administration Press; 1980.
- 42 Donabedian A. The quality of care: How can it be assessed? *JAMA* 1988 23-30 sept;260(12):1743–1748.
- 43 Abernethy MA, Horne M, Lillis AM, Malina MA, Selto FH. A multi-method approach to building causal performance maps from expert knowledge. *Management Accounting Research* 2005 Jun;16(2):135–155.
- 44 Adair CE, Simpson L, Birdsell JM, Omelchuk K, Casebeer AL, Gerdiner HP, et al. *Performance measurement systems in health and mental health services: Models, practices and effectiveness*. A State of the Science Review; Submitted to: The Alberta Heritage Foundation for Medical Research. 2003.
- 45 Adair CE, Simpson E, Casebeer AL, Birdsell JM, Hayden KA, Lewis S. Performance measurement in healthcare: Part 1 - Concepts and trends from a state of the science review. *Healthc Policy* 2006 May;1(4):85–104.
- 46 Adair CE, Simpson E, Casebeer AL, Birdsell JM, Hayden KA, Lewis S. Performance measurement in healthcare: Part II - State of the science findings by stage of the performance. *Healthc Policy* 2006 Jul;2(1):56–78.
- 47 Arah OA, Klazinga NS, Delnoij DMJ, Ten Asbroek AHA, Custers T. Conceptual frameworks for health systems performance: A quest for effectiveness, quality, and improvement. *Int J Qual Health Care* 2003 Oct;15(3):377–398.
- 48 Baker GR, MacIntosh-Murray A, Porcellato C, Dionne L, Stelmacovich K, Born K. *High performing healthcare systems: Delivering quality by design*. Toronto, Ontario: Longwoods Publishing; 2008.
- 49 Baker GR, Brooks N, Anderson G, Brown A, McKillop I, Murray M, et al. Healthcare performance measurement in Canada: Who's doing what? *Healthc Q* 1998;2(2):22–26.
- 50 Hurst J. Performance measurement and improvement in OECD health systems: Overview of issues and challenges. In: Organisation of Economic Co-operation and Development, ed. *Measuring up: Improving health system performance in OECD countries*. Paris, France: Organisation of Economic Co-operation and Development; 2002. 35–54.
- 51 Hurst J, Jee-Hughes M. *Performance measurement and performance management in OECD health systems*. Organisation for Economic Co-operation and Development, Labour Market and Social Policy Occasional Papers No. 47, January. 2001. [http://www.oilis.oecd.org/OLIS/2000DOC.NSF/LINKTO/DEELSA-ELSA-WD\(2000\)8](http://www.oilis.oecd.org/OLIS/2000DOC.NSF/LINKTO/DEELSA-ELSA-WD(2000)8).

- 52 Kelley E, Hurst J. *Health care quality indicators project: Conceptual framework paper*. OECD Health Working Papers No. 23. Paris, France: Organisation for Economic Co-operation and Development, 9 March. 2006.
<http://www.oecd.org/dataoecd/1/36/36262363.pdf>.
- 53 Kelley E, Hurst J. *Health care quality indicators project: Initial indicators report*. OECD Health Working Papers No. 22. Paris, France: Organisation for Economic Co-operation and Development, 9 March. 2006.
<http://www.oecd.org/dataoecd/1/34/36262514.pdf>.
- 54 Smith P. Performance management in British health care: Will it deliver? *Health Aff* 2002;21(3):103–115.
- 55 Phillips CD, Chen M, Sherman M. To what degree does provider performance affect a quality indicator? The case of nursing homes and ADL change. *Gerontologist* 2008;48(3):330–337.
- 56 Barnsley J, Lemieux-Charles L, Baker R. Selecting clinical outcome indicators for monitoring quality of care. *Healthc Manage Forum* 1996 Spring;9(1):5–21.
- 57 Linder JA, Kaleba EO, Kmetik KS. Using electronic health records to measure physician performance for acute conditions in primary care: Empirical evaluation of the community-acquired pneumonia clinical quality measure set. *Med Care* 2009 Feb;47(2):208–216.
- 58 Loeb JM. The current state of performance measurement in health care. *Int J Qual Health Care* 2004;16(1):i5–i9.
- 59 Klazinga N, Stronks K, Delnoij D, Verhoeff A. Indicators without a cause: Reflections on the development and use of indicators in health care from a public health perspective. *Int J Qual Health Care* 2001 Dec;13(6):433–438.
- 60 Lebas MJ. Performance measurement and performance management. *International Journal of Production Economics* 1995;41(1-3):23–35.
- 61 Folan P, Browne J. A review of performance measurement: Towards performance management. *Computers in Industry* 2005 Sep;56(7):663–680.
- 62 Ammons DN. Overcoming the inadequacies of performance measurement in local government: The case of libraries and leisure services. *Public Admin R* 1995 Jan/Feb;55(1):37–47.
- 63 Boucckaert G. Measurement and meaningful management. *Public Productivity and Management Review* 1993 Fall;17(1):31–43.
- 64 Julnes P de L. 1st ed. *Performance-based management systems*. Boca Raton, FL: CRC Press; 2008.

- 65 Poister TH, Streib G. Performance measurement in municipal government: Assessing the state of the practice. *Public Admin R* 1999 Jul/Aug;59(4):325–335.
- 66 Kuhlmann S. Performance measurement in European local governments: A comparative analysis of reform experiences in Great Britain, France, Sweden and Germany. *International Review of Administrative Sciences* 2010 Jul 6;76(2):331–345.
- 67 Chassin MR, Galvin RW, National Roundtable on Health Care Quality. The urgent need to improve health care quality: Institute of Medicine National roundtable on Health Care Quality. *JAMA* 1998 Sep 16;280(11):1000–1005.
- 68 McGlynn EA, Asch SM. Developing a clinical performance measure. *Am J Prev Med* 1998 Apr;14(sup. 3):14–21.
- 69 Baker GR, Norton PG, Flintoft V, Blais R, Brown AD, Cox J, et al. The Canadian adverse events study: The incidence of adverse events among hospital patients in Canada. *Can Med Assoc J* 2004 May 25;170(11):1678–1686.
- 70 Myers R, Lacey R. Consumer satisfaction, performance and accountability in the public sector. *International Review of Administrative Sciences* 1996 3 Quarter;62(3):331–350.
- 71 Monahan PJ. *Chaoulli v Quebec and the future of Canadian healthcare: Patient accountability as the "sixth principle" of the Canada Health Act*. Benefactors Lecture, Toronto, ON: C.D. Howe Institute, ISPCO Inc., 29 November. 2006. http://www.cdhowe.org/pdf/benefactors_lecture_2006.pdf.
- 72 Behn R. 1st ed. *Rethinking democratic accountability*. Washington DC: Brookings Institution Press; 2001.
- 73 Deber R, Topp A, Zakus D. *Private delivery and public goals: Mechanisms for ensuring that hospitals meet public objectives*. World Bank, Background Paper, May 22. 2004. <http://siteresources.worldbank.org/INTHSD/Resources/376278-1202320704235/GuidingPrivHospitalsDeberetal.pdf>.
- 74 Townley B. *Critical views of performance measurement*. Encyclopedia of Social Measurement, Vol. 1. 2005.
- 75 ten Asbroek AH, Arah OA, Geelhoed J, Custers T, Delnoij DM, Klazinga NS. Developing a national performance indicator framework for the Dutch health system. *Int J Qual Health Care* 2004 Apr;16(sup. 1):i65–i75.
- 76 Baker RG, Pink GH. A balanced scorecard for Canadian hospitals. *Healthc Manage Forum* 1995 Winter;8(4):7–13.
- 77 Pink GH, McKillop I, Schraa EG, Preyra C, Montgomery C, Baker GR. Creating a balanced scorecard for a hospital system. *Journal of Health Care Finance* 2001 Spring;27(3):1–20.

- 78 Kaplan RS, Norton DP. The balanced scorecard: Measures that drive performance. *Harv Bus Rev* 1992 Jan/Feb.
- 79 Lindsay P, Schull M, Bronskill S, Anderson G. The development of indicators to measure the quality of clinical care in emergency departments following a modified-Delphi approach. *Academic Emergency Medicine* 2002 Nov;9(11):1131–1139.
- 80 Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Br Med J* 2002;326:16–19.
- 81 Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD health care quality indicators project. *Int J Qual Health Care* 2006 Sep;18(sup. 1):5–13.
- 82 Armstrong R, Waters E, Moore L, Riggs E, Cuervo LG, Lumbiganon P, et al. Improving the reporting of public health intervention research: Advancing TREND and CONSORT. *Journal of Public Health* 2008 Mar;30(1):103–109.
- 83 Tregunno D, Baker R, Barnsley J, Murray M. Competing values of emergency department performance: Balancing multiple stakeholder perspectives. *Health Serv Res* 2004 Aug;39(4).
- 84 Hamblin R. Regulation, measurements and incentives. The experience in the US and the UK: Does context matter? *Journal of the Royal Society for the Promotion of Health* 2008 Nov;128(6):291–298.
- 85 Irwin R. Managing performance: An introduction. *Eurohealth* 2010;16(3):15–16.
- 86 Exworthy M. The performance paradigm in the English NHS: Potential, pitfalls, and prospects. *Eurohealth* 2010;16(3):16–19.
- 87 Bevan G. If neither altruism nor markets have improved NHS performance, what might? *Eurohealth* 2010;16(3):20–22.
- 88 Ham C. Improving performance in the English National Health Service. *Eurohealth* 2010;16(3):23–25.
- 89 Klazinga N. Health system performance management. *Eurohealth* 2010;16(3):26–28.
- 90 Provincial Auditor of Ontario. *Annual Report of the Office of the Provincial Auditor of Ontario*. Toronto: Office of the Provincial Auditor of Ontario. 2003. <http://www.auditor.on.ca/en/reports2003en.htm>.
- 91 Propper C, Sutton M, Whitnall C, Windmeijer F. *Did 'targets and terror' reduce waiting times in England for hospital care?* Working Paper No. 07/179, Bristol, UK: Centre for Market and Public Organisation, Bristol Institute of Public Affairs, University of Bristol, November. 2007. Downloaded from <http://www.bris.ac.uk/Depts/CMPO/workingpapers/wp179pdf>.

- 92 Alexander JA, Weiner BJ, Shortell SM, Baker LC, Becker MP. The role of organizational infrastructure in implementation of hospitals' quality improvement. *Hosp Top* 2006 Winter;84(1):11–20.
- 93 Aitken J-M. Voices from the inside: Managing district health services in Nepal. *Int J Health Plann Manage* 1994 Oct-Dec;9(4):309–340.
- 94 Lipsky M. 1st ed. *Street-level bureaucracy: Dilemmas of the individual in public services*. New York: Russell-Sage Foundation Publications; 1980.
- 95 Leggat SG, Narine L, Lemieux-Charles L, Barnsley J, Baker GR, Sicotte C, et al. A review of organizational performance assessment in health care. *Health Serv Manage Res* 1998;11(1):3–18.
- 96 Howlett M, Ramesh M. Patterns of policy instrument choice: Policy styles, policy learning and the privatization experience. *Policy Studies Review* 1993 Spring/Summer;12(1/2):3–24.
- 97 Howlett M. Managing the "hollow state": Procedural policy instruments and modern governance. *Can Pub Adm* 2000 Winter;43(4):412–431.
- 98 Howlett M, Ramesh M. 2nd ed. *Studying public policy: Policy cycles and policy subsystems*. New York, NY: Oxford University Press; 2003.
- 99 Doern GB, Phidd RW. 2nd ed. *Canadian public policy: Ideas, structure, process*. Toronto, ON: Nelson Canada; 1992.
- 100 Eliadis FP. *Instrument choice*. Policy Research Initiative - Foundation Paper, Ottawa, ON: Government of Canada, updated 27 February. 2007. http://policyresearch.gc.ca/page.asp?pagenm=law-droit_instrument_foundation.
- 101 Baxter-Moore N. Policy implementation and the role of the state: A revised approach to the study of policy instruments. In: Jackson RJ, Jackson D, Baxter-Moore N, eds. *Contemporary Canadian politics: Reading and notes*. Scarborough, Ontario: Prentice-Hall Canada Inc.; 1987. 336–355.
- 102 Evans RG. *Strained mercy: The economics of Canadian health care*. Toronto, ON: Butterworths & Co.; 1984.
- 103 Robinson JC. Theory and practice in the design of physician payment incentives. *Milbank Q* 2001 May;79(2):149–177.
- 104 Wodchis WP, Fries BE, Pollack H. Payer incentives and physical rehabilitation therapy for nonelderly institutional long-term care residents: Evidence from Michigan and Ontario. *Arch Phys Med Rehabil* 2004 Feb;85(2):210–217.
- 105 Donaldson C, Gerard K, Jan S, Mitton C, Wiseman V. 2nd ed. *Economics of health care financing: The visible hand*. Basingstoke: Palgrave/Macmillan; 2005.

- 106 Walshe K. 1st ed. *Regulating Healthcare: A prescription for improvement?* Maidenhead: McGraw Hill; 2003.
- 107 Eisenhardt KM. Agency theory: An assessment and review. *Acad Mgt R* 1989 Jan;14(1):57–74.
- 108 McGuire TG. Physician agency. In: Culyer AJ, Newhouse JP, eds. *Handbook of health economics*. Amsterdam, The Netherlands: Elsevier; 2000. 461–536.
- 109 Saltman RB, Busse R, Mossailos E. *Regulating entrepreneurial behaviour in European healthcare systems*. Buckingham: Open University Press; 2002.
- 110 Chinitz D, Preker A, Wasem J. Balancing competition and solidarity in health care financing. In: Saltman RB, Figueras J, Sakellarides C, eds. *Critical Challenges for Health Care Reform in Europe*. Buckingham: Open University Press; 1998. 55–77.
- 111 Allegrante JP, Barry MM, Auld ME, Lamarre M-C, Taub A. Toward international collaboration on credentialing in health promotion and health education: The Galway consensus conference. *Health Education & Behavior* 2009 Jun;36(3):427–438.
- 112 Baker EL, Stevens RH. Linking agency accreditation to workforce credentialing: A few steps along a difficult path. *J Public Health Manag Pract* 2007 Jul/Aug;13(4):430–431.
- 113 Ayer TS. Accreditation through standards of excellence for public health organizations. *J Public Health Manag Pract* 1998 Jul;4(4):24–27.
- 114 Howells G. The potential and limits of consumer empowerment by information. *Journal of Law and Society* 2005 Sep;32(3):349–370.
- 115 Morris K, Zelmer J. *Public reporting of performance measures in health care*. Health Care Accountability Papers, No. 4, Health Network, Ottawa, ON: Canadian Policy Research Networks, February. 2005. <http://www.cprn.com/doc.cfm?doc=1176>.
- 116 Barnsley J, Berta W, Cockerill R, MacPhail J, Vayda E. Identifying performance indicators for family practice: Assessing levels of consensus. *Can Fam Physician* 2005 May;51:700–701.
- 117 Hospital Report Research Collaborative. *Hospital report 2001: Acute care*. Ottawa: Canadian Institute for Health Information. 2001. <https://secure.cihi.ca/estore/productFamily.htm?pf=PFC323&lang=en&media=0>.
- 118 Organisation for Economic Co-operation and Development. *Measuring up: Improving health system performance in OECD countries*. Paris, France: Organization for Economic Co-operation and Development. 2002. <http://www.ikwilwerken.nl/pdf/eu/8102011healthsystem.pdf>.

- 119 Shaw C. *What are the methods of hospital performance measurement?* Copenhagen, Denmark: World Health Organization Regional Office for Europe. 2003. <http://www.who.dk/HEN/Syntheses/hospper/200308223>.
- 120 Veillard J, Champagne F, Klazinga N, Kazandjian V, Arah OA, Guisset A-L. A performance assessment framework for hospitals: The WHO regional office for Europe PATH project. *Int J Qual Health Care* 2005 Dec;17(6):487–496.
- 121 Saltman RB, Ferroussier-Davis O. The concept of stewardship in health policy. *Bull World Health Organ* 2000;78(6):732–739.
- 122 Lemieux-Charles L, Champagne F. 1st ed. *Using knowledge and evidence in health care: Multidisciplinary perspectives*. Toronto, ON: University of Toronto Press; 2004.
- 123 Aylward S, Stolee P, Keat N, Johncox V. Effectiveness of continuing education in long-term care: A literature review. *Gerontologist* 2003 K;42(2):259–271.
- 124 Abelson J, Guavin F-P. *Engaging citizens: One route to health care accountability*. Health Care Accountability Papers, No. 2, Health Network, Ottawa, ON: Canadian Policy Research Networks, April. 2004. <http://www.cprn.com/doc.cfm?doc=560>.
- 125 Ries NM, Caulfield T. *Accountability in health care and legal approaches*. Health Care Accountability Papers, No. 3, Health Network, Ottawa, ON: Canadian Policy Research Networks, May. 2004. <http://www.cprn.com/doc.cfm?doc=622>.
- 126 Shah CP. 5th ed. *Public health and preventive medicine in Canada*. Toronto: Elsevier Canada; 2003.
- 127 Last JM. 2nd ed. *A dictionary of epidemiology*. New York: Oxford University Press; 1988.
- 128 Health Disparities Task Group of the Federal/Provincial/Territorial Advisory Committee on Population Health and Health Security. *Reducing health disparities - Roles of the health sector: Discussion paper*. December 2004. 2004. <http://www.phac-aspc.gc.ca/ph-sp/disparities/ddp-eng.php>.
- 129 British Columbia Ministry of Health. *A framework for core functions in public health*. Population Health and Wellness. Resource document, March. 2005. http://www.phabc.org/pdfcore/core_functions.pdf?NSNST_Flood=hrkhhdyo.
- 130 Moloughney BW. *The renewal of public health in Nova Scotia: Building a public health system to meet the needs of Nova Scotians*. Final Report, Halifax: Government of Nova Scotia. 2006. http://www.gov.ns.ca/heal/publichealth/content/pubs/07148_TheRenewalOfHinNSReport_Apr06En.pdf.

- 131 Public Health Research and Knowledge Translation Network. *Background documents for public health research priorities in Atlantic Canada*. Halifax: Atlantic Health Promotion Research Centre, Dalhousie University, steering committee co-chairs: Lynn McIntyre, Janet Braunstein Moody, and Donna Meagher-Stewart, December. 2005.
<http://preventionresearch.dal.ca/Web%20Version%20PH.pdf>.
- 132 National Public Health Partnership. *National Delphi study on public health functions in Australia: Report on the findings*. Melbourne, Victoria, Australia, January. 2000.
www.dhs.vic.gov.au/nphp/publications/phpractice/delphi-body.pdf.
- 133 Bettcher DW, Sapirie S, Goon EH. Essential public health functions: Results of the international Delphi study. *World Health Stat Q* 1998;51(1):44–54.
- 134 Canada. Parliament. Standing Senate Committee on Social Affairs Science and Technology. *Reforming health protection and promotion in Canada: Time to act*. Report, November. 2003. <http://www.parl.gc.ca/37/2/parlbus/commbus/senate/com-e/soci-e/repfinnov03-e.htm>.
- 135 Jorm L, Gruszyn S, Churches T. A multidimensional classification of public health activity in Australia. *Aust New Zealand Health Policy* 2009 Apr 9;6(9):1–10.
- 136 Capacity Review Committee. *Revitalizing Ontario's public health capacity: A discussion of issues and opinions. Final report*. Interim report, November. 2005.
http://www.health.gov.on.ca/english/public/pub/ministry_reports/capacity_review06/capacity_review06.pdf.
- 137 Starfield Consulting Ltd. *Capacity review committee, phase II stakeholder consultation: Accountabilities, funding and governance report*. January 12. 2006.
http://www.health.gov.on.ca/english/public/pub/ministry_reports/capacity_review06/capacity_review06.html.
- 138 Capacity Review Committee. *Revitalizing Ontario's public health capacity: The final report of the Capacity Review Committee*. Final Report, Ontario Ministry of Health and Long-Term Care, Toronto, ON: May. 2006. Downloaded from
http://www.health.gov.on.ca/english/public/pub/ministry_reports/capacity_review06/capacityreview06.pdf.
- 139 Ministry of Health and Long-Term Care. *Ontario public health standards*. (2010).
<http://www.health.gov.on.ca/english/providers/program/pubhealth/ophstandards/ophs/ophprotocols.html>.
- 140 Deber R, McDougall C, Wilson K. *Financing and delivering public health in Canada's provinces and territories: Approaches, issues, and options*. Report prepared for Public Health Agency of Canada: Toronto, Ontario, March 30. 2006.
- 141 Government of Ontario. *Ontario health pandemic influenza plan*. June. 2005.
<http://www.ontla.on.ca/library/repository/mon/8000/244106.pdf>.

- 142 National Advisory Committee on SARS and Public Health. *Learning from SARS: Renewal of public health in Canada*. Ottawa, ON: Health Canada, Naylor Report, October. 2003. <http://www.phac-aspc.gc.ca/publicat/sars-sras/naylor/>.
- 143 Expert Panel on SARS and Infectious Disease Control. *Interim report of the Ontario expert panel*. Toronto: Ministry of Health and Long-Term Care, 12 December. 2003. http://www.health.gov.on.ca/english/public/pub/ministry_reports/walker_panel_2003/introduction.pdf.
- 144 Expert Panel on SARS and Infectious Disease Control. *For the public's health: A plan of action. Final report of the Ontario expert panel on SARS and infectious disease control*. Toronto: Ministry of Health and Long-Term Care, April. 2004. http://www.health.gov.on.ca/english/public/pub/ministry_reports/walker04/walker04_mn.html.
- 145 Campbell A. *SARS and public health in Ontario. The SARS commission interim report*. Toronto: Ministry of Health and Long-Term Care, 15 April. 2004. <http://www.health.gov.on.ca/english/public/pub/ministryreports/campbell04/campbell04.html>.
- 146 Basrur S, Yaffe B, Henry B. SARS: A local public health perspective. *Can J Public Health* 2004 Jan/Feb;95(1):22–24.
- 147 D'Cunha C. SARS: Lessons learned from a provincial perspective. *Can J Public Health* 2004 Jan/Feb;95(1):25–26.
- 148 Campbell A. *The SARS Commission: Second interim report: SARS and public health legislation*. Toronto: Ministry of Health and Long-Term Care, 5 April. 2005. <http://www.health.gov.on.ca/english/public/pub/ministryreports/campbell04/campbell04.html>.
- 149 Basrur SV. *Building the foundation of a strong public health system for Ontarians*. Annual report of the Chief Medical Officer of Health to the Ontario Legislative Assembly. 2005. http://www.health.gov.on.ca/english/public/pub/ministry_reports/ph_ontario/ph_ontario.html.
- 150 Van Kersbergen K, Van Waarden F. 'Governance' as a bridge between disciplines: Cross-disciplinary inspiration regarding shifts in governance and problems of governability, accountability and legitimacy. *Eur J Polit Res* 2004 Mar;43(2):143–171.
- 151 Kempa M, Shearing C, Burriss S. *Changes in governance: A background review*. Prepared for the Salzburg Seminar on the Governance of Health. 2005. http://www.temple.edu/lawschool/phrhcs/salzburg/Global_health_Governance_Review.pdf.
- 152 Bertelli A. The role of political ideology in the structural design of new governance agencies. *Public Admin R* 2006 Jul/Aug;66(4):583–595.

- 153 Jordan A, Wurzel RKW, Zito A. The rise of 'new' policy instruments in comparative perspective: Has governance eclipsed government? *Political Studies* 2005 Oct;53(3):477–496.
- 154 Denis J-L. Governance and management of change in Canada's health system. In: Forest P-G, Marchildon GP, McIntosh T, eds. *Changing health care in Canada: Romanow papers, Volume 2*. Toronto: University of Toronto Press; 2004. 82–114.
- 155 Denis J-L, Champagne F, Pomey M-P, Prével J, Tré G. *Toward a framework for the analysis of governance in health care organizations and systems*. Preliminary report presented to the Canadian Council on Health Services Accreditation, Ottawa University of Montréal, October. 2005.
- 156 Osborne DE, Gaebler T. 1st ed. *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. Reading, Massachusetts: Addison-Wesley Publications; 1992.
- 157 Ricketts TC. How many physicians? How much does it matter? *JAMA* 2009 Oct 21;302(15):1701–1702.
- 158 Horwitz J. Nonprofit ownership, private property, and public accountability. *Health Aff* 2006 Jun 20;25:w308–w311.
- 159 Deber R. Delivering health care services: Public, not-for-profit, or private? In: Marchildon GP, McIntosh T, Forest P-G, eds. *The fiscal sustainability of health care in Canada: Romanow Papers, Volume 1*. Toronto, ON: University of Toronto Press; 2004. 233–296.
- 160 Sloan FA. Not-for-profit ownership and hospital behavior. In: Culyer AJ, Newhouse JP, eds. *Handbook of health economics*. Amsterdam, Netherlands: Elsevier; 2000. 1141–1174.
- 161 Docteur E, Oxley H. *Health-care systems: Lessons from the reform experience*. OECD Health Working Papers, No. 9, OECD Publishing. 2003.
<http://www.oecd.org/LongAbstract/0,2546,en26493458721547787119684111,00.html>.
- 162 Schwartz R. Evaluating government-Third sector collaboration: Political considerations. In: Gray A, Jenkins B, Mayne J, Leeuw F, eds. *Exploring Evaluation and Collaborative Government*. New Brunswick,NJ: Transaction; 2003. 83–104.
- 163 Schwartz R. The elusiveness of accountability in public management. In: Vigoda-Gadot E, ed. *Public administration - The new generation: An interdisciplinary critical analysis*. Kentucky: Routledge, USA; 2002. 63–84.
- 164 Schwartz R. Managing government - Third sector collaboration: Accountability, ambiguity and politics. *Int J Pub Admin* 2001 Aug;24(11):1161–1188.

- 165 Schwartz R, Sharkansky I. Collaboration with the 'Third Sector' - Issues of accountability: Mapping Israeli versions of this problematic. *Public Policy and Administration* 2000 Fall;15(2):92–106.
- 166 Rhodes RAW. *Understanding governance: Policy networks, governance, reflexivity and accountability*. Buckingham, UK: Open University Press; 1997.
- 167 Preker AS, Harding A. 1st ed. *Innovations in health service delivery: The corporatization of public hospitals*. Washington, DC: The World Bank; 2003.
- 168 Brandsen T, Boogers M, Tops P. Soft governance, hard consequences: The ambiguous status of unofficial guidelines. *Public Admin R* 2006 Jul/Aug;66(4):546–553.
- 169 Thomas PG. From good intentions to successful implementation: The case of patient safety in Canada. *Can Pub Adm* 2006 Winter;49(4):415–441.
- 170 Baker R. *Contribution of governance to patient safety initiatives in Australia, England, New Zealand and the United States*. Health Policy Research Program, Summary of Research Results, Toronto, ON: University of Toronto, January. 2005. <http://www.hc-sc.gc.ca/sr-sr/pubs/hpr-rpms/2005-baker-eng.php>.
- 171 Preker AS, Xingzhu L, Velenyi EV. 1st ed. *Public ends, private means: Strategic purchasing of health services*. Washington DC: World Bank Publications; 2007.
- 172 Harding A, Preker AS. 1st ed. *Private participation in health services*. Washington: The World Bank; 2003.
- 173 Vining AR, Globerman S. Contracting-out health care services: A conceptual framework. *Health Policy* 1999 Jan;46(2):77–96.
- 174 Preker AS, Harding A, Travis P. "Make or buy" decisions in the production of health care goods and services: new insights from institutional economics and organizational theory. *Bull World Health Organ* 2000 Jun;78(6):779–789.
- 175 Rico A, Puig-Junoy J. What can we learn from the regulation of public utilities? In: Saltman RB, Busse R, Mossialos E, eds. *Regulating Entrepreneurial Behaviour in European Health Care Systems*. Buckingham: Open University Press; 2002. 73–90.
- 176 Preker AS, Harding A. *The economics of public and private roles in health care: Insights from institutional economics and organizational theory*. Health, Nutrition and Population Discussion Paper, Human Development Network, World Bank, June. 2000. <http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/Preker-TheEconomicsOfPublic-whole.pdf>.

- 177 Jakab M, Preker A, Harding A, Hawkins L. *The introduction of market forces in the public hospital sector: From new public sector management to organizational reform*. Washington, DC: World Bank's Human Development Network, Health, Nutrition, and Population, June. 2002.
<http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resource/s/281627-1095698140167/Jakab-TheIntroductionof-whole.pdf>.
- 178 Williamson OE. The economics of organization: The transaction cost approach. *Am J Soc* 1981 Nov;87(3):548–577.
- 179 Williamson OE. Public and private bureaucracies: A transaction cost economics perspective. *Journal of Law, Economics & Organization* 1999 Mar 1;15(1):306–342.
- 180 Williamson OE. 1st ed. *The economic institutions of capitalism*. New York, NY: The Free Press; 1985.
- 181 Peckham S. The new general practice contract and reform of primary care in the United Kingdom. *Healthc Policy* 2007 May;2(4):34–48.
- 182 Peckham S, Hann A. General practice and public health: Assessing the impact of the new GMS contract. *Critical Public Health* 2008 Sep;18(3):347–356.
- 183 Nolte E, Bain C, McKee M. Population health. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, MA: Cambridge University Press; 2009. 27–62.
- 184 McGlynn EA. Measuring clinical quality and appropriateness. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, MA: Cambridge University Press; 2009. 87–113.
- 185 Smith PC, Busse R. Targets and performance measurement. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, MA: Cambridge University Press; 2009. 509–536.
- 186 Saskatchewan Health and Regional Health Services. *Performance management accountability indicators*. Regional Policy Branch, Planning and Reporting, The Regional Health Services Act, Sections 54(2)(a) and 55, June. 2007.
<http://www.health.gov.sk.ca/performance-accountability-06/07>.
- 187 Conrad DA. Incentives for health-care performance improvement. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, MA: Cambridge University Press; 2009. 582–612.

- 188 Sequist TD, Bates DW. Developing information technology capacity for performance measurement. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge, MA: Cambridge University Press; 2009. 552–581.
- 189 Simonet D. The new public management theory and European health-care reforms. *Can Pub Adm* 2008 Dec;51(4):617–635.
- 190 Freeman T. Using performance indicators to improve health care quality in the public sector: A review of the literature. *Health Serv Manage Res* 2002 May;15(2):126–137.