The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1751-1879.htm

Evaluation of a leadership development impact assessment toolkit: a comparative case study of experts' perspectives in three Canadian provinces

Received 13 June 2022
Revised 30 August 2022
24 October 2022
Accepted 31 October 2022

Mehri Karimi-Dehkordi
Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada

Graham Dickson
Professor Emeritus, Royal Roads University, Victoria, Canada, and Canadian Society of Physician Leaders, British Columbia, Canada

Kelly Grimes
Canadian Health Leadership Network, Ottawa, Canada

Suzanne Schell
ROI Institute Canada, Ottawa, Canada, and

Ivy Bourgeault
Telfer School of Management, University of Ottawa, Ottawa, Canada

Abstract

Purpose – This paper aims to explore users' perceptions of whether the Leadership Development Impact Assessment (LDI) Toolkit is valid, reliable, simple to use and cost-effective as a guide to its quality improvement.

Design/methodology/approach – The Canadian Health Leadership Network codesigned and codeveloped the LDI Toolkit as a theory-driven and evidence-informed resource that aims to assist health-care organizational development practitioners to evaluate various programs at five levels of impact: reaction, learning, application, impact and return on investment (ROI) and intangible benefits. A comparative evaluative case study was conducted using online questionnaires and semistructured telephone interviews with three health organizations where robust leadership development programs were in place. A total of seven leadership consultants and specialists participated from three Canadian provinces. Data were analyzed sequentially in two stages involving descriptive statistical analysis augmented with a qualitative content analysis of key themes.

Findings – Users perceived the toolkit as cost-effective in terms of direct costs, indirect costs and intangibles; they found it easy-to-use in terms of clarity, logic and structure, ease of navigation with a coherent layout; and they assessed the sources of the evidence-informed tools and guides as appropriate.

The authors want to thank leadership consultants and specialists as study participants from British Columbia, Saskatchewan, and Nova Scotia for their continuous and long-term contributions to this research.

Funding: This study was supported by funding from a Mitacs Accelerate Award No. IT10345 in partnership with the Canadian Health Leadership Network (CHLNet), both awarded to Mehri Karimi-Dehkordi, a postdoctoral fellow with the University of Ottawa.

Leadership in Health Services
© Emerald Publishing Limited
1751-1879
DOI 10.1108/LHS-06-2022-0068

Users rated the toolkit highly on their perceptions of its validity and reliability. The analysis also informed the refinement of the toolkit.

Originality/value – The refined LDI Toolkit is a comprehensive online collection of various tools to support health organizations to evaluate the leadership development investments effectively and efficiently at five impact levels including ROI.

Keywords Leadership development, Toolkit, User experiences, Evaluation

Paper type Research paper

Introduction

Effective leadership is essential in promoting a positive culture, enhancing the quality of care, improving patients' outcomes and ensuring the success of health-care organizations (Sfantou *et al.*, 2017; Dickson and Tholl, 2020). To this end, health-care organizations have invested in leadership development (LD) initiatives and programs to build capacity and effectiveness at the individual, team and organizational levels (Dickson and Van Aerde, 2018).

LD programs and initiatives used in different health-care settings vary in terms of the length, cost and objectives. As a result, the evaluation of the programs was frequently reported as complex and challenging often failing to connect to key outcomes (Phillips *et al.*, 2015a). Specifically, the literature highlights the difficulty of measuring the impact of complex interventions (Streatfield and Markless, 2019; Davys *et al.*, 2017; Dave *et al.*, 2021; Moldoveanu and Narayandas, 2019; Hopkins, and Meyer, 2019) and on moving beyond process-type of evaluation (Dave *et al.*, 2021; Davys *et al.*, 2017). Challenges include but are not limited to the lack of evaluation skills and resources for organizations (Davys *et al.*, 2017), a paucity of rigorous evaluations of LD programs to use as comparisons (Joseph-Richard and McCray, 2022) and difficulty in quantifying the return on investment (ROI) for programs (Day *et al.*, 2021). Also critical is the need for more comprehensive multimethod, multilevel, systematic and robust evaluations including a focus on outcome and impact (Collins and Denyer, 2008; Dave *et al.*, 2021; Jeyaraman *et al.*, 2018; Newstead *et al.*, 2020; Njah *et al.*, 2021; Wallace *et al.*, 2021). This ongoing challenge was also stressed in a King's Fund review, which highlighted the lack of literature on the effectiveness of the LD programs and called for the need for evidence-based approaches to ensure ROI (West *et al.*, 2015).

To address these gaps, the Canadian Health Leadership Network (CHLNet), a coalition of 40+ organizations (called Network Partners) who gather to build health leadership capacity and capabilities across Canada, developed a toolkit of resources, tools and guidance to facilitate organizational efforts to determine the impact and effectiveness of a wide variety of LD initiatives at different levels of evaluation.

CHLNet leadership development impact toolkit

The development of the LDI toolkit

In 2015, CHLNet initiated a project to design and develop a centralized LDI Toolkit. In the process of CHLNet's consultations and presentations to its network partners [1], it becomes clear that the ROI in LD is neither known nor clear for Canadian health-care organizations. Without such data, investment in LD programs is difficult to procure. Thus, the CHLNet's LDI Toolkit was developed to adhere to the following criteria: evidence-based, simple to use, customized to different clients and programs and cost-effective. Because most organizations invest in LD to improve individual leadership capability and build collective leadership capacity in an organization aimed at creating healthy workplace cultures (Dickson and

Tholl, 2020), the toolkit focused on leadership both at the individual (leader) development as well as organizational (leadership) development, something that is often not done in LD evaluations (Wallace *et al.*, 2021).

The toolkit was developed in three stages. *First*, the Research and Evaluation group took on the role of a steering group. They appointed project team leads and the leads then engaged other experts from member partner organizations as required. A deep dive into the literature on leadership development, evaluation of education programs, and methodologies relative to determining the ROI of leadership development programs was undertaken to inform the framing of the project (https://chlnet.ca/ldi-toolkit-2). Second, the Research and Evaluation group commissioned a study undertaken by the Centre for Health Innovation at the University of Manitoba to "scope out" the dimensions of such a project (Jeyaraman *et al.*, 2018). This work helped the project team to design a schema for the toolkit that became the focus of its development. Third, the CHLNet project team partnered with the ROI Institute of Canada, which provided advice, guidance and additional materials that were used to augment resources. Toolkit was developed using the Phillips ROI Methodology™: Levels of data and guiding principles (Phillips *et al.*, 2015a, 2015b). The toolkit was designed for online access for member partners only.

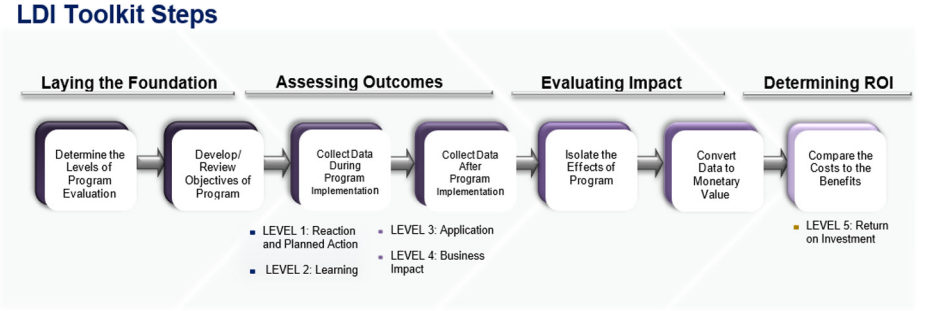
The content of the LDI toolkit

The LDI Toolkit illustrates four main steps, including laying the foundation, assessing outcomes, evaluating impact and determining ROI to elucidate the evaluation process of LD programs. Each step includes guidelines, templates, tools, downloadable sources and exemplars of the implications of tools. Figure 1 outlines the LDI Toolkit steps with the tools, which are organized under each step. The examples of the instruments include the validated and reliable tools (e.g. ISA Engagement Scale and Organizational Empowerment Scale) and the tools (e.g. the ROI Institute's Level of Data and Evaluation and Methods to Convert Impact Measures to Monetary Values) adapted from the ROI Institute Canada, which have been field tested and validated (ROI Institute Canada, 2022). Features which differentiate the LDI Toolkit from other existing toolkits and resources are: its logical format and navigating web pages based with concise information rather than a document with information overload; its user-friendly focus including instructions, tools and examples for all levels of impact evaluation levels specific to health care; and supporting users by the availability of consultation with the ROI Methodology specialist for complex LD programs. Figure 1 provides an overview of the components of the toolkit.

The evaluation of the LDI toolkit

The toolkit evaluation and refinement was an iterative process that occurred through two main pathways. The evaluation of the toolkit began from the start of the LDI Toolkit development in an evolving and enduring manner. Once the toolkit was designed, CHLNet subsequently commissioned a formative evaluation of its utility – i.e. did it meet the criteria established by the project team? To do so, the toolkit was evaluated by expert users, which is the focus of this study. Prior to users' evaluation, the toolkit was reviewed by the Research and Evaluation Working Group comprised of field experts using a consensus approach to assess content validity (Crawford and Kelder, 2019), and then it became available online on the CHLNet website with the involvement of a web designer. The evaluation of the LDI aimed to:

- evaluate the properties of the LDI Toolkit to determine whether the toolkit is perceived elegantly simple, valid, reliable and cost-effective; and
- guide the refinement of the toolkit.



The LDI Toolkit illustrates four main steps to elucidate the evaluation process of LD programs. Each step includes downloadable sources, guidelines, templates, tools, and examples of implications of tools:

- **Laying the Foundation**, which addresses how to determine the levels of program evaluation and develop or review the program objectives.

Downloadable Files

- Program Alignment Developing Objectives – Example and How to Use
- Levels of Data and Evaluation
- A Conversation about Evaluating Programs at the Different Levels
- Criteria for Levels of Evaluation

- **Assessing Outcomes**, which focuses on the necessity of determining the types of personal and organizational metrics that need to be measured. Predetermined desired outcomes inform the selection of the optimal tools in advance.

Downloadable Files

- Sample Questionnaire Follow Up – Sample Questionnaire
- Reaction Questionnaire – Sample Questionnaire
- Self Sassement Application Questionnaire – Sample Questionnaire
- Questionnaire Design Steps
- Action Plan – Example and How to Use
- Data Collection Tools – Guidelines
- Data Collection Plan
- Sample Data Collection Tools and Online Tools

- **Evaluating Impact**, which addresses isolating the tangible and intangible benefits that are directly from the program.

Downloadable Files

- Isolating the Benefits of the Leadership Development – Isolating the Impact of the Program
- 5 Steps to Convert Impact Measures to Monetary Value
- Methods to Convert Impact Measures to Monetary Values
- Leadership Development – Examples of Data – Impact Measures Quantitative and Qualitative

- **Determining ROI**, which refers to the fifth level of evaluation, which is an ultimate measure and a profitability ratio for evaluating the efficiency of a program.

Downloadable Files

- The 10 Steps in the ROI Methodology
- Capturing Program Costs / Fully Loaded Cost Worksheet
- Evidence Brief: Development of a Return on Investment Tool for Healthcare Leadership Development
- Article: Return on investment in healthcare leadership development programs

Figure 1.
Overview of leadership development impact toolkit components

To assess whether the toolkit goals were achieved, the assessment focused on five levels of data collected during the implementation of a program, including 1) reaction, 2) learning, 3) application, 4) impact and 5) ROI and intangible benefits (Phillips *et al.*, 2015a; Phillips *et al.*, 2015b; ROI Institute Canada, 2022).

Methodology

A multimethod comparative evaluative case study of user experiences was conducted using self-administered questionnaires and semistructured interviews (Aberdeen, 2013; Fletcher *et al.*, 2015)

in three health organizations across three Canadian provinces where robust leadership development programs were in place. Users' and stakeholders' perspectives have been taken into account in the literature to address the ease of use (Connell *et al.*, 2021), cost-effectiveness (Jones Rhodes *et al.*, 2018; Feldman *et al.*, 2020), validity (Connell *et al.*, 2021; Lusiana *et al.*, 2011) and reliability components (Eliasson *et al.*, 2021; Møller and Wiberg, 2020).

To be included in the study, pilot sites: could be any association representing health professionals in a province, or any provincial health services authority in Canada that:

- planned to make substantial investments into LD programs between 2017 and 2018;
- aimed to evaluate the impact of the LD programs; and
- volunteered to use the toolkit to measure outcomes and impacts of their LD programs and provide feedback.

Site and program information

Three sites in three provinces in Canada: Nova Scotia, Saskatchewan and British Columbia, met the criteria for selection. In the pilot sites, for each respective target group, as noted below, varied interventions were used (see online supplementary Table S1) via different delivery strategies (i.e. classroom-based sessions, individual and group coaching sessions, e-learning, triad learning, action learning projects and online collaborative virtual platforms). Class size and mix of participants (health professionals, level of leader) varied widely among the sites and included:

- 24 physicians with different specialties and different backgrounds from different regions;
- 59 existing and aspiring managers, directors, physician leaders and program team leads who sought experiential learning and stretch opportunities to enhance personal leadership and maximize their potential within teams; and
- 48 managers who were within the first 18 months of their initial appointments.

Data collection

Data collection was carried out in two phases in the Spring and Fall of 2018. A total of seven leadership consultants and specialists, who had master's level degrees and additional leadership/management related certifications and engaged with LD programs planning and evaluation, were involved with the collection of data across the two phases after their virtual orientation sessions and using the online toolkit.

Phase 1: The first-round data collection was conducted between April and May 2018; data were collected using a questionnaire administered via telephone interviews, taking up to 90 min.

To establish face validity and content validity, the questionnaire was developed based on a literature review and eight expert ROI steering group member consultation and pretesting. The questionnaire included 40 Likert-scaled (10-point) [2] open-ended questions and subquestions that were divided into five main sections:

(1) characteristics of the site and program in which the toolkit were used.;
(2) perspectives on the validity and reliability of the toolkit;
(3) perspectives on the ease of use of the toolkit (i.e. clarity, logic and time required to understand the toolkit and if an assistant was required);

(4) views on the cost-effectiveness, including direct costs, indirect costs [3] and intangible benefits, [4] compared with: the cost of hiring a consultant and the time that they would have spent to find, to learn and to use alternative tools; and

(5) overall evaluation and feedback.

Phase 2: The second round of mixed methods data collection was carried out in Fall 2018. The same participants in one site were available to take part in Phase 2, whereas participants of the two other sites differed from the first round. A modified version of the Phase 1 questionnaire was used to ensure that the questions are measuring the properties of the toolkit over time and applicability when it was used to assess the LD programs at higher impact levels.

Data analysis and synthesis

Data were first analyzed based on demographic information regarding each site and its programs, and with attention to the objectives of the study: perspectives on validity, reliability, ease of use and cost-effectiveness of the LDI Toolkit. Likert data were analyzed descriptively using mean values, frequency, percentage and cumulative frequency distribution of the measures gathered in the surveys. Data were graphically depicted across the three participating sites for interpretation.

The qualitative analysis was performed through an iterative content analysis process. To this end the interviews were audio recorded and transcribed verbatim. The coding process began after the first interview was transcribed. Upon completion of coding for all transcripts, we then developed themes within and across the cases. Two types of themes were involved, emerging (e.g. all in one place, well-organized online format) and priori themes (e.g. time invested in understanding the toolkit, challenges users experienced). We then created overarching themes for each perceived quality dimension of the toolkit and areas that required further improvement.

The results of each phase were distributed among the ROI Steering Group team. Frequent meetings were held with team members to discuss the results and make effective and collective plans through brainstorming to refine the toolkit and its evaluation.

Ethical considerations

Participants voluntarily took part in this study and provided consent for the audio recording before the commencement of the data collection. Data collection was conducted at the time per suggestion by leadership consultants. Participants' quotes are presented anonymously and only aggregated quantitative data are reported to ensure confidentiality. Data related to the LD programs are presented with no identity information. In consultation with the University of Ottawa Research Ethics Board, it was deemed that a review was not required given the evaluative nature of the research.

Findings

Project participants

Figure 2 shows an equal number of participants in the two phases of this study who identify as men or as women in the LD programs in one site but a higher number of women in the other two sites than men. One participant recognized this imbalance as its recruitment pool contained a high number of nurses, who themselves disproportionately identify as women. Regardless of varying distribution, all participants stated that gender was not a criterion for participation in the study.

Toolkit features used

In general, the organizations invested in LD programs to improve individual leadership capability and collective leadership capacity in the organization. Table 1 shows the tools chosen by the pilot sites to evaluate the LD programs. The pilot sites chose to use the toolkit to measure the impact of the program specifically to:

- measure the effectiveness of the components of the program, to be able to continue in subsequent years;
- make continuous improvements in the program content/delivery;
- forward results to the stakeholders; and
- help their organization develop.

Analysis of two rounds and types of data collected address two overarching areas: the evaluation of the toolkit and the aspects of the toolkit requiring further refinement identified through low scores on the Likert-scale data and the qualitative data that provide specific suggestions for improvement.

Evaluation of the LDI toolkit

Decisions to use the LDI Toolkit were influenced by two main attributes: its comprehensiveness and the quality of the tools it encompassed.

The perceived comprehensiveness of the toolkit. The toolkit provides guidance helping the user to assess the process of evaluating the programs effectively and systematically in that it:

- guides users in identifying different types of information (e.g. indicators, intangibles);
- guides users in constructing their customized measurement plan and evaluation for the used programs;
- contains a guide to help users to identify, select and use the most appropriate tools;
- provides users options to find, choose and use questions that they may need to develop a tailored approach;
- equips users to collect data throughout the program and after completion of the program;
- encourages leaders to get involved with the program; and
- presents tools that could be used for each desired impact level, including level 1 (reaction and planned action), level 2 (learning and confidence), level 3 (application and implementation), level 4 (business impact) and level 5 (ROI), and intangibles.

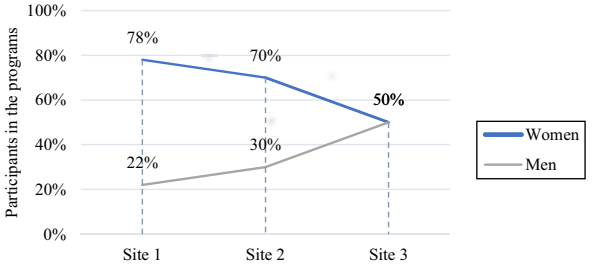


Figure 2.
Gender breakdown of participants across three pilot program sites

Level of evaluation	Examples of measures selected
Level 1 Reaction and planned action	• Overall satisfaction with each learning module • Overall satisfaction with the Leadership Program • Reaction toward the training program • The value of the action plan
Level 2 Learning	• Learning new knowledge/skills in each learning module
Level 3 Application	• Taking action toward their personal learning goals, including: – Ability to use knowledge/skills from each learning module – Percentage of participants who applied knowledge/skills – Ability to use coaching skills – Ability to use conflict management skills – Ability to lead effective teams – Ability to use systems thinking concepts – Extent to which participants were equipped to lead, support and engage others in change in the context of their work • Completing a meaningful change throughout the program • Implementing action plan
Level 4 Impact	• Achieving effective change leadership practices organizationally • Increasing personal leadership capacity in their workday through dealing with conflict proactively • Increasing engagement and problem-solving in huddles and interactions with staff • Creating capacity in their workday
Level 5 ROI	• Calculating ROI

Note: *ROI Institute Canada (2022)

Table 1. Example of measures used by the sites based on five evaluation levels of Phillips ROI Methodology™*

While the pilot sites collectively intended to evaluate all the impact levels, the sites individually varied in their preferred impact levels to measure. Organizations most desired to measure level 4 (all three sites) followed by level 5 (two of the sites).

The outcomes showed that despite the primary desires, the total number of tools selected by all the sites collectively were most for level 3 (application), then followed equally for level 1 (reaction) and level 4 (impact). These tools consisted of the reaction questionnaire, self-assessment application, action plan, level of the data evaluation framework, developing objectives, follow-up questionnaire, reaction questionnaire, leadership development questionnaire and the program alignment tool.

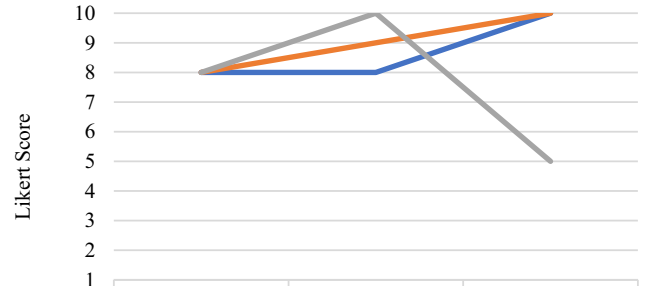
Overall, participants felt that the tools they had selected to evaluate their complex programs, including learning modules, mentorship, coaching, individualized programs failed to evaluate the desired levels of impact in particular level 4 impact. The perceived

benefits were also extended to its potential ability to speed up the evaluation process and allow the pilot sites to identify which aspects of the program require enhancement.

The perceived quality dimensions of the toolkit. The quality of the toolkit was appraised on the four dimensions including users' perceptions on its validity, reliability, ease of use and cost-effectiveness. Users reported that the LDI Toolkit was effective in the process of program evaluation at various levels.

Perceptions of the validity of the toolkit. In this context, user perspectives about the validity of the toolkit were explored subjectively on whether the evaluation tools are sound concerning true value, construct soundness and content appropriateness (Figure 3):

- *Perspectives on true value* refer to the extent that toolkit users agree that the tools appear to measure the LD program outcomes and impacts. All participants consistently ranked an 8 on the 10-point Likert scale. Pilot sites qualified their rating by stating that the toolkit appears to measure LD program outcomes and impacts due to the toolkit layout, which contains ample content, extensive rich examples and extensive questionnaires within each level.
- *Perspectives on construct soundness* refer to the extent that toolkit users agree that the tools measure what it claims to measure. A mean score of 9 on the 10-point Likert scale was identified. Participants liked that there are various data collection options in the toolkit, which allowed users to choose an appropriate tool for a specific context; however, participants identified challenges related to assessing changes in leadership capability and converting capability data to monetary data.
- *Perspectives on content appropriateness* refer to the extent that which toolkit users agree that the tools are appropriate, considering the aim of the toolkit. While one site scored a 5, two other sites scored a 10 on the Likert scale because the toolkit contains the tools to measure impacts and outcomes of the LD programs, develops a customized measure for the evaluation of LD programs and met their specific needs for certain tools.



	Perspectives on true value	Perspectives on construct soundness	Perspectives on content appropriateness
Site 1	8	8	10
Site 2	8	9	10
Site 3	8	10	5

Note: Likert score ranges from 1 (very low) to 10 (very high)

Figure 3.
User perspectives on true value, construct soundness and content appropriateness of the LDI Toolkit across three sites

Perceptions of the reliability of the toolkit. Perceptions on reliability refer to the extent to which toolkit users agree about the reproducibility of a measure in the same population. Participants commented that the findings were expected to be accurate as tools are evidence-based and described the toolkit as a source of reproducible tools. All participants across all sites consistently reported they would use the same tools in the same way to evaluate another similar program. Two aspects were further particularly explored:

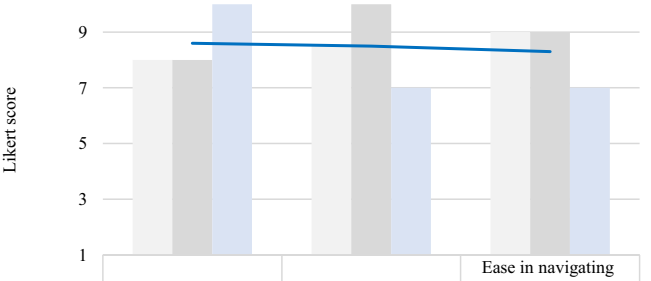
(1) *Perceptions on consistency* refer to the extent that the toolkit was perceived in terms of the interrelatedness of the items in the toolkit. The mean of the participants' scores on the internal consistency of the toolkit on the Likert scale was 7.6.

(2) *Perceptions on responsiveness* refer to the extent that the toolkit was perceived to be useful for monitoring changes in outcomes over time. The mean score was 8 on the Likert scale. Participants of two sites rated and perceived that the tools are sensitive to change; however, in one site, the participant could not answer this question because they had not measured anything more than once.

Perceptions of the ease of use of the toolkit. All participants consistently responded with an 8 for their perceptions about the toolkit as a simple, elegant and user-friendly resource. As noted earlier, the respondents' perspectives on ease of use of the toolkit were further assessed based on three subcriteria, including clarity, logic/structure and ease of navigation. As Figure 4 shows, the highest mean score on the Likert scale was 8.6 for the *clarity* of the toolkit. The mean score of the two other criteria, *logic/structure* and *navigation*, was 8.3.

Analysis of the qualitative data revealed two overarching criteria of the ease of use of the toolkit: (1) its structure and function and (2) its contents. Participants overtly expressed five key attributes related to the *structure* of the toolkit, which makes it user friendly: (1) the clear layout of the design, (2) the logic of the interface, (3) the flow of the components, (4) the use of hyperlinks and (5) the labels for documents/tools. As one user in British Columbia stated:

> The Toolkit overview on the right-hand column of the page makes navigation quite easy to go out and back to different parts of the Toolkit. The Downloadable Evaluation Resources at the bottom



	Clarity	Logic and structure	Ease in navigating through the website/ accessibility to tools
Site 1	8	8.5	9
Site 2	8	10	9
Site 3	10	7	7
Mean	8.6	8.5	8.3

Note: Likert score ranges from 1 (very low) to 10 (very high)

Figure 4.
User perspectives on clarity, logic and ease of use of the LDI Toolkit across three sites

of the overview section is great to see all the links on one page without having to go in and out to get to them.

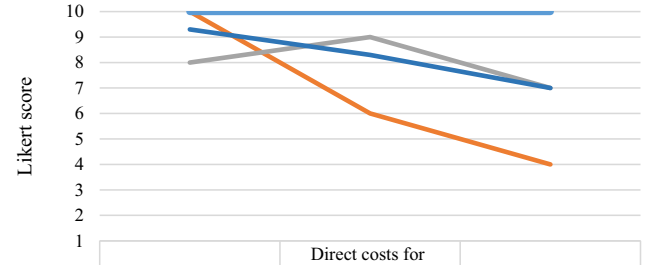
The *content* of the LDI Toolkit makes it simple and elegant and helps users in determining needed measures or questions. It includes:

- a variety of tools and perception of toolkit as a one-stop shop for all resources;
- the succinct description;
- the straightforward/simple content;
- the clear description/introduction for each section;
- the helpful information;
- the comprehensive content; and
- the general guidance.

Perceptions of the cost-effectiveness of the toolkit. In general, the cost-effectiveness analysis was reflected in the participants' general perceptions, in addition to specific cost-effectiveness parameters: direct costs, indirect or productivity costs, as well as intangible factors. (Figure 5).

Participants' general perceptions. In general, participants perceived the toolkit as worth the effort, as they only needed to invest a couple of hours to understand the toolkit; and it provides tools that help them save time. The mean number of hours that were invested for the understanding of the toolkit was 4.3 h. All participants stated that this amount of time was not excessive. As one respondent stated: "So, putting a couple of hours upfront wouldn't be excessive." The mean hours provided for assistance was 1.2 h. Assistance was provided by the ROI expert from the ROI Institute Canada and was fourfold: to offer users advice and feedback, to help them in thinking through the toolkit, to help them to identify and develop customized measures for different levels and to provide them with examples.

Perceptions on direct costs referred to participants' views on the time that was used to learn and to practice the LDI Toolkit, as opposed to the cost of hiring a consultant firm to do



	Direct time costs	Direct costs for time: The LDI Toolkit vs alternatives	Indirect time costs
Site 1	10	6	4
Site 2	8	9	7
Site 3	10	10	10
Mean	9.3	8.3	7

Note: Likert score ranges from 1 (very low) to 10 (very high)

Figure 5.
User perspectives on cost-effectiveness analysis of the LDI Toolkit across three sites

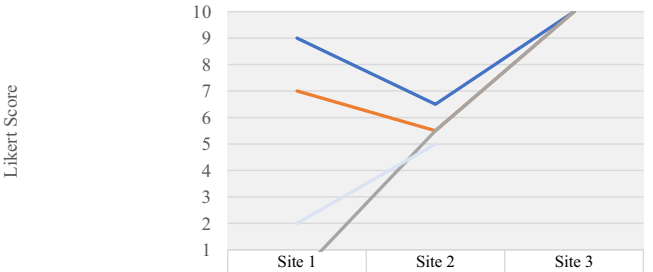
the work. When comparing the direct cost of time used to learn and to use the toolkit versus the cost of hiring a consultant firm to do the work, the mean of the Likert scores on the direct cost-effectiveness of the LDI Toolkit was 9.3. The major reasons for their rating on the direct cost-effectiveness of the LDI Toolkit were as follows: The toolkit:

- provides a comprehensive suite of tools;
- eliminates the need to hire a consultant;
- includes an element of efficiency; and
- allows users to gain familiarity with it over time, which may potentially improve the effectiveness of the toolkit.

In addition, the mean of the Likert scores of respondents' opinion on direct costs for the time used to learn and use the LDI Toolkit versus the time to find, understand and use alternative tools from other sources was 8.3. Participants' responses included: "the toolkit saved time, as all tools were available in one place."

Perceptions on indirect costs, as another aspect of cost-effectiveness, refer to views on the extent that learning and using the toolkit influenced the level of productivity as compared to the time they used their traditional approaches to find out the tailored tools to evaluate the used LD program. It was ranked as a 10, a 7 and a 4, with a mean of 7 by the three pilot sites. It was noted that increased productivity is due to the simplicity and the ease of use of the toolkit, requiring less time to learn. Participants in two sites, however, qualified their ratings with references to the time that was invested in learning the LDI Toolkit, as it interfered with other areas of their work, and it temporarily slowed down the LD program process.

Perceptions on intangibles, as another parameter of cost-effectiveness analysis, refers to views on the potential benefits that are associated with the use of the LDI Toolkit, including the level of satisfaction, level of stress, a sense of organizational commitment and level of teamwork as compared to their subjective feelings prior to using the toolkit to evaluate their used LD programs [5] (Figure 6). Of these four intangibles, the mean of Likert ranking scores for the level of satisfaction, level of stress, a sense of organizational commitment and level of teamwork were 8.5, 5.2, 5.2 and 3.5, respectively. With a mean of 8.5, the level of satisfaction was the highest mean level of intangible benefits using the toolkit. The perceived



	Site 1	Site 2	Site 3
Level of satisfaction	9	6.5	10
Level of stress	7	5.5	10
A sense of organizational commitment	0	5.5	10
Level of teamwork	2	5	

Note: Likert score ranges from 1 (very low) to 10 (very high)

Figure 6.
User perspectives on the intangible benefits of using the LDI Toolkit across three sites

satisfaction was frequently linked to the Toolkit attribute as "one-stop to all of the downloadable resources." In general, self-related benefits (the level of satisfaction and stress) were more dominant than the organization-related benefits (level of teamwork and a sense of organizational commitment).

Despite the assessment that the toolkit was found to be logical and easy to follow overall, some areas were perceived as requiring further clarifications, and in other areas deemed as too detailed or theoretical, requiring simplification.

Refinement of the LDI toolkit

The refinement process was calibrated with the second objective of the case study, which aimed to guide the toolkit improvements in such that it perceived more simple, valid, reliable and cost-effective. In general, we used three strategies during the refinement process: enhancing the simplicity of the toolkit, refining the users' guide and creating a service registry.

Increasing simplicity. Respondents' feedback suggested that the toolkit can be made more straightforward and more user-friendly in at least four ways. First, more simple language could facilitate toolkit use; reducing the amount of theoretical information could help make it easier to understand. Second, content mapping of the LDI Toolkit can give users a visual overview of the LDI Toolkit. Third, reducing the number of steps (links) could expedite reaching the desired tools. Fourth, greater consistency in language could help users to navigate efficiently through the toolkit. To achieve some of these goals, we both simplified the language, offered further examples, tools and explanation of the key aspects of the content. A diagram was also created to illustrate a big picture along with the pertinent information about the evaluation process,

Refining the user's guide. Respondents' feedback suggested that to improve the usefulness and ease of use of the toolkit, the user's guide should clearly explain what the tool is and how to use it. To improve the users' guide, further instructions and examples were added (for instance, for converting data to monetary data). A frequently asked question section was also created to address comments and technical questions that were not related to the structures and content of the LDI Toolkit. This section is also intended to serve as a quick reference to provide guidance for users. A virtual and instructive session was added to the toolkit, covering three main sections:

(1) the main steps taken in the development process (i.e. scoping review, toolkit design, the results of the case study and the refinement process);

(2) the four main stages of evaluation included in the toolkit; and

(3) the two types of supportive approaches available for the users.

Creating a registry of users and a repository of projects. Users also suggested creating a supportive approach to promote and facilitate collaboration, learning and sharing of knowledge and experiences. To address deficiencies observed, a registry and repository of projects were designed to help organizations seeking mentor/mentee relationships by providing insight and guidance regarding evaluating their leadership improvement programs. Such learning opportunities may effectively and efficiently accelerate the evaluation process in an organization (Conlon, 2004).

Discussion

This comparative case study was motivated by the need to examine users' perception of the ease of use, and cost-effectiveness of a novel LDI Toolkit. Essentially, the LDI Toolkit is a

theory-driven and evidence-informed source that aims to assist health-care organizational development practitioners to evaluate various LD programs at five levels of impact. The process of toolkit development and evaluation was aligned with the suggestions provided in a systematic review (Yamada *et al.*, 2015) that toolkits should be constructed based on theory and evidence and rigorously evaluated to identify factors contributing to their effectiveness and uptake (Yamada *et al.*, 2015). This study provides promising results in support of the LDI Toolkit. The toolkit was used by leadership consultants and specialists from three organizations that used varied LD programs. Users needed tools to measure the effectiveness, outcomes and impacts of the multifaceted initiatives among target groups with varying proportions of gender distribution. Experts provided their perspectives on the quality dimension of the toolkit and areas that required further improvement. Notably, participants reported that the toolkit provided the guidance and tools to evaluate their programs effectively and systematically. The full effectiveness of the toolkit emerged as the ability of the toolkit to evaluate the LD programs systematically and efficiently at multiple levels and to determine which area requires improvement subsequently. This finding is similar to previous studies wherein effectiveness was discussed as a criterion for success; and as a means to: provide tools tailored to the needs of the users, facilitate effective use of evidence across a variety of health-care settings and enable users to identify challenges surrounding the provision of care (Dunne *et al.*, 2015; Davis *et al.*, 2017, Yamada *et al.*, 2015).

At face value, users also rated the toolkit highly for their perceptions on the true value, construct soundness, content appropriateness and responsiveness and described the toolkit as an accommodating and appropriate source of evidence-based tools and guides with a coherent layout. The need for valid tools for leadership development assessment was presented as a substantive gap in a recent systematic review (Crawford and Kelder, 2019). Our study, thus, adds evidence to address this gap. We argue that the quality aspects are fundamentally related to rigorous methods that underpin the toolkit development by means of drawing on the empirical evidence (Jeyaraman *et al.*, 2018), the LEADS in a Caring Environment leadership capabilities framework (Dickson and Van Aerde, 2018) and the ROI Methodology (Phillips *et al.*, 2015a; Phillips *et al.*, 2015b; ROI Institute Canada (2022) and the collaboration of field experts.

The ease of use of the toolkit was highly rated and corresponded to the clarity, logic and ease of navigation. Data also explicitly referred to content and the structure of the toolkit. Sridhar *et al.* (2021) revealed that ease of use is a determinant for toolkit uptake, and its absence was described as a salient barrier to toolkit use. This study adds to and supports the rising body of evidence on the characteristics that make a tool easy to use and the likelihood of acceptance of the toolkit in health-care organizations.

Users perceived cost-effectiveness as also an essential characteristic of the toolkit. The benefits corresponded to reducing the direct costs (owing to the consultant firm expenses and time costs), indirect costs (owing to improving productivity level) and intangibles (owing to increasing the level of satisfaction). Cost-effectiveness has been explicitly addressed in the context of the program evaluation process; the evaluation costs can contribute to draining resources from the program itself (Killion, 2008). As toolkit was best rated for being efficient by using fewer resources (i.e. time, consultant) in program evaluation, the findings could be of interest to policymakers in health-care organizations. However, the impact of using the toolkit on the sense of organizational commitment and level of teamwork was rated low by one site. We expect the evaluation of the refined toolkit may lead to a different view.

Overall, the toolkit was supported by users for the above three qualities. In line with these attributes' merits, a recent systematic review showed that barriers might hinder

toolkit uptake consisted of but not limited to time restrictions, costs, difficulty to use and dissatisfaction with the toolkit's content (Hempel *et al.*, 2019). The iterative refinement process suggests that the continuous supports tailored to organizations' needs likely increases a toolkit's uptake (Mudge *et al.*, 2020).

Originality/practical implications

The evaluation and recommendations identified from this study informed the refinement process of the LDI Toolkit. The refinement process thus offers insights that can inform the design and development of future toolkits. This study can be used as evidence to introduce the LDI Toolkit to those seeking a comprehensive toolkit to evaluate LD programs supported by users in terms of their perceptions on the ease of use, cost-effectiveness, validity and sensitivity to change. The toolkit is available online and adapted for health-care organizations that use various LD programs based on their structures, priorities and end goals and desire to assess different impact levels (Dickson and Tholl, 2020).

Limitations and areas for future research

Participants in this study were credible leadership consultants involved in planning, implementing, monitoring and evaluating the various LD programs in health-related organizations to strengthen the data's corroboration and tackle potential biases. Nevertheless, the current study was subject to some methodological limitations.

First, there are limits to the generalizability of the findings due to the small sample size, and insight was not obtained in international contexts. Second, there was an unavoidable concern regarding the level of program evaluation that the pilot sites had carried out; only in one site, the ROI measurement was underway after the effects of the LD program had been isolated. The third limitation is related to the finding of quality aspects of the toolkit reflects how these characteristics appeared to experts who used the Toolkit. While the subjective views of the toolkit users were crucial in understanding the toolkit attributes and the areas which required refinement, objective measurement of psychometric properties of the toolkit may yield different results. These limitations suggest new areas for future research in health-related organizations. Research is thus warranted to assess the psychometric properties of the toolkit using a statistical analysis approach. Cost-effectiveness analyses in health-care organizations can also be designed to assign a monetary value to the outcomes using the toolkit. Future research can tease out the merits of the LDI Toolkit in supporting generalization within the broader health care context by involving a wide target audience at the national and international levels.

Notes

1. CHLNet is a social enterprise of 40+ organizations called "network partners" who gather around health leadership founded on the twin principles of trust and reciprocity. Members cut across jurisdictions, policymakers, academics, health associations, regional health authorities, patients and health disciplines. Part of our work is to reduce duplication and share emerging leadership practices for 21st-century care. In addition to many opportunities to dialogue in engage, network partners, meet twice a year at semi-annual network partner roundtables. Visit www.chlnet.ca to learn more.
2. A higher value on a 10-point Likert scale (ranging from 1 to 10) represented a higher degree of participant's agreement with the statement.

3. Indirect costs were referred to as the sense of productivity and the respondents' perspectives of the time they spent and the extent to which the process of learning and using the toolkit influenced the level of their sense of productivity.
4. Intangible benefits were referred to as their level of satisfaction, level of stress, a sense of organizational commitment and level of teamworking.
5. The corresponding question was:

   Please rate, on a scale of 1–10, the extents of intangible benefits of using the Toolkit, outlined below, with (1) indicating the lowest level and (10) indicating the highest level. Please provide specific comments/examples.

   Level of your satisfaction

   Level of your stress

   A sense of organizational commitment

   Level of teamworking

References

Aberdeen, T. (2013), "Yin, RK (2009), case study research: design and methods. Thousand oaks, CA: Sage", *The Canadian Journal of Action Research*, Vol. 14 No. 1, pp. 69-71.

Collins, J. and Denyer, D. (2008), "Leadership learning and development: a framework for evaluation", in James K.T. and Collins J. (Eds), *Leadership Learning*, Palgrave Macmillan, London, pp. 161-177.

Conlon, T.J. (2004), "A review of informal learning literature, theory and implications for practice in developing global professional competence", *Journal of European Industrial Training*, Vol. 28 Nos 2/3/4, pp. 283-295.

Connell, S.A., Winkelmann, Z.K. and Games, K.E. (2021), "Athletic trainers' exposure to telemedicine influences perspectives and intention to use", *Athletic Training and Sports Health Care*, Vol. 13 No. 4, pp. e193-e201.

Crawford, J.A. and Kelder, J.A. (2019), "Do we measure leadership effectively? Articulating and evaluating scale development psychometrics for best practice", *The Leadership Quarterly*, Vol. 30 No. 1, pp. 133-144.

Dave, G., Noble, C., Chandler, C., Corbie-Smith, G. and Fernandez, C.S. (2021), "Clinical scholars: using program evaluation to inform leadership development", in Fernandez, C.S.P. and Corbie-Smith, G. (Eds), *Leading Community Based Changes in the Culture of Health in the US – Experiences in Developing the Team and Impacting the Community*, IntechOpen, London.

Davis, M.M., Howk, S., Spurlock, M., McGinnis, P.B., Cohen, D.J. and Fagnan, L.J. (2017), "A qualitative study of clinic and community member perspectives on intervention toolkits: unless the toolkit is used it won't help solve the problem", *BMC Health Services Research*, Vol. 17 No. 1, pp. 1-9.

Davys, A.M., O'Connell, M., May, J. and Burns, B. (2017), "Evaluation of professional supervision in Aotearoa/New Zealand: an interprofessional study", *International Journal of Mental Health Nursing*, Vol. 26 No. 3, pp. 249-258.

Day, D.V., Riggio, R.E., Tan, S.J. and Conger, J.A. (2021), "Advancing the science of 21st-century leadership development: theory, research, and practice", *The Leadership Quarterly*, Vol. 32 No. 5, p. 101557.

Dickson, G. and Tholl, B. (2020), *Bringing Leadership to Life in Health: LEADS in a Caring Environment: Putting LEADS to Work*, Springer-Verlag, London.

Dickson, G. and Van Aerde, J. (2018), "Enabling physicians to lead: Canada's LEADS framework", *Leadership in Health Services*, Vol. 31 No. 2, pp. 183-194.

Dunne, S., Lunn, C., Kirwan, M., Matthews, A. and Condell, S. (2015), "Planning and selecting evaluation designs for leadership training: a toolkit for nurse managers and educators", *Journal of Professional Nursing*, Vol. 31 No. 6, pp. 475-481.

Eliasson, K., Dahlgren, G., Hellman, T., Lewis, C., Palm, P., Svartengren, M. and Nyman, T. (2021), "Company representatives' experiences of occupational health surveillance for workers exposed to hand-intensive work: a qualitative study", *International Journal of Environmental Research and Public Health*, Vol. 18 No. 4, pp. 1-18, doi: 10.3390/ijerph18042018.

Feldman, R.M., Cioffi, G.A., Liebmann, J.M. and Weinreb, R.N. (2020), "Current knowledge and attitudes concerning cost-effectiveness in glaucoma pharmacotherapy: a glaucoma specialists focus group study", *Clinical Ophthalmology*, Vol. 14, pp. 729-739.

Fletcher, A.J., MacPhee, M. and Dickson, G. (2015), "Doing participatory action research in a multicase study: a methodological example", *International Journal of Qualitative Methods*, Vol. 14 No. 5, pp. 1-9.

Hempel, S., O'Hanlon, C., Lim, Y.W., Danz, M., Larkin, J. and Rubenstein, L. (2019), "Spread tools: a systematic review of components, uptake, and effectiveness of quality improvement toolkits", *Implementation Science*, Vol. 14 No. 1, pp. 1-46.

Hopkins, K. and Meyer, M. (2019), "Evaluating behavioral and organizational outcomes of leadership development in human service organizations", *Human Service Organizations Management Leadership and Governance*, Vol. 43 No. 4, pp. 290-298.

Jeyaraman, M.M., Qadar, S.M.Z., Wierzbowski, A., Farshidfar, F., Lys, J., Dickson, G., Grimes, K., Phillips, L.A., Mitchell, J.I., Van Aerde, J. and Johnson, D. (2018), "Return on investment in healthcare leadership development programs", *Leadership in Health Services*, Vol. 31 No. 1, pp. 77-97.

Jones Rhodes, W.C., Ritzwoller, D.P. and Glasgow, R.E. (2018), "Stakeholder perspectives on costs and resource expenditures: tools for addressing economic issues most relevant to patients, providers, and clinics", *Translational Behavioral Medicine*, Vol. 8 No. 5, pp. 675-682.

Joseph-Richard, P. and McCray, J. (2022), "Evaluating leadership development in a changing world? Alternative models and approaches for healthcare organisations", *Human Resource Development International*, pp. 1-37.

Killion, J. (2008), *Assessing Impact: Evaluating Staff Development*, Corwin Press, Thousand Oaks, CA.

Lusiana, B., van Noordwijk, M., Suyamto, D., Mulia, R., Joshi, L. and Cadisch, G. (2011), "Users' perspectives on validity of a simulation model for natural resource management", *International Journal of Agricultural Sustainability*, Vol. 9 No. 2, pp. 364-378.

Moldoveanu, M. and Narayandas, D. (2019), "The future of leadership development", *Harvard Business Review*, Vol. 97 No. 2, pp. 40-48.

Møller, A. and Wiberg, S. (2020), "A qualitative approach to examining how digital twins impact users' mental models", Master's thesis, University of Bergen, Noway.

Mudge, S., Sezier, A., Payne, D., Smith, G. and Kayes, N. (2020), "Pilot trial of the living well toolkit: qualitative analysis and implications for refinement and future implementation", *BMC Health Services Research*, Vol. 20 No. 1, pp. 1-12.

Newstead, T., Dawkins, S., Macklin, R. and Martin, A. (2020), "Evaluating the virtues project as a leadership development programme", *Leadership*, Vol. 16 No. 6, pp. 633-660.

Njah, J., Hansoti, B., Adeyami, A., Bruce, K., O'Malley, G., Gugerty, M.K., Chi, B.H., Lubimbi, N., Steen, E., Stampfly, S., Berman, E. and Kimball, A.M. (2021), "Measuring for success: evaluating leadership training programs for sustainable impact", *Annals of Global Health*, Vol. 87 No. 1, pp. 1-10.

Phillips, P.P., Phillips, J.J. and Ray, R. (2015a), *Measuring the Success of Leadership Development: A Step-by-Step Guide for Measuring Impact and Calculating ROI*, Association for Talent Development, Alexandria, VA.

Phillips, J.J., Phillips, P.P. and Schell, S.C. (2015b), "The value of motivation: how to measure the value, impact, and ROI of motivational projects, programs, and events", *Performance Improvement*, Vol. 54 No. 5, pp. 16-27.

ROI Institute Canada (2022), "ROI methodology", *Retrieved August 2018*, available at: https://roiinstitutecanada.com/roi-methodology

Sfantou, D.F., Laliotis, A., Patelarou, A.E., Sifaki-Pistolla, D., Matalliotakis, M. and Patelarou, E. (2017), "Importance of leadership style towards quality of care measures in healthcare settings: a systematic review", *Healthcare*, Vol. 5 No. 4, pp. 1-17.

Sridhar, A., Drahota, A. and Walsworth, K. (2021), "Facilitators and barriers to the utilization of the ACT SMART implementation toolkit in community-based organizations: a qualitative study", *Implementation Science Communications*, Vol. 2 No. 1, pp. 1-14.

Streatfield, D. and Markless, S. (2019), "Impact evaluation and IFLA: evaluating the impact of three international capacity building initiatives", *Performance Measurement and Metrics*, Vol. 20 No. 2, pp. 105-122.

Wallace, D.M., Torres, E.M. and Zaccaro, S.J. (2021), "Just what do we think we are doing? Learning outcomes of leader and leadership development", *The Leadership Quarterly*, Vol. 32 No. 5, p. 101494.

West, M., Armit, K., Loewenthal, L., Eckert, R., West, T. and Lee, A. (2015), *Leadership and Leadership Development in Healthcare: The Evidence Base*, The King's Fund, London.

Yamada, J., Shorkey, A., Barwick, M., Widger, K. and Stevens, B.J. (2015), "The effectiveness of toolkits as knowledge translation strategies for integrating evidence into clinical care: a systematic review", *BMJ Open*, Vol. 5 No. 4, p. e006808.

Further reading

Canadian Health Leadership Network: CHLNet (2020), "Leadership development impact assessment toolkit (LDI)", available at: https://chlnet.ca

Supplementary material

The supplementary material for this article can be found online.

About the authors

Mehri Karimi-Dehkordi is a postdoctoral fellow at the University of Alberta. She was involved in this pan-Canadian research as a postdoctoral fellow of the University of Ottawa, following her PhD in Nursing at the University of Alberta. Mehri has developed a modified form of critical interpretive synthesis to synthesize qualitative and quantitative data. She has made many contributions to mixed methods and qualitative research. Her contributions have also extended to the development of several resources that aim to improve the quality of care. Her areas of research interest include leadership development evaluation, quality care improvement with an emphasis on the quality indicators and person-centered care. Mehri Karimi-Dehkordi is the corresponding author and can be contacted at: karimide@ualberta.ca

Graham Dickson is Professor Emeritus at Royal Roads University in Canada. While at RRU, Graham helped to develop the Master of Arts in Leadership for Royal Roads University and was the founding Director of the Centre for Health Leadership and Research. Currently he is working as the Senior Research Advisor to the Canadian Society of Physician Leaders and the Canadian Health Leadership Network. He is a member of the LEADS Collaborative, an enterprise established to develop leadership within the health sector in Canada. Graham was principal investigator in a cross-Canada research project on LEADS in a Caring Environment capabilities framework. LEADS is now endorsed by health authorities in 10 provinces, and by most national health organizations.

Kelly Grimes is CHLNet's Executive Director and responsible for its strategic and business plans for this value network in consultation with its 42 Network Partners. She has participated in various health leadership studies, most recently the Return on Investment in Leadership Development, Leading Through COVID Action Research study, Wise Practices of Leadership Development and the Canadian Health Leadership Benchmarking studies (I and II). Kelly has been an author or contributor on numerous health policy reports, which have assisted federal, provincial and local governments in their health policy decision-making.

Suzanne Schell, CRP is the CEO of ROI Institute Canada. Suzanne has consulted and taught the ROI Methodology™ to numerous health-care professionals and organizations across Canada. She has implemented Return on Investment evaluations for programs and improvement initiatives in all areas of health care. Suzanne was a codeveloper of the evaluation component of the LDI, incorporating the ROI methodology. Suzanne has authored many ROI reports which have resulted in successful funding for programs and improvement initiatives.

Ivy Bourgeault, PhD, is a Professor in the Telfer School of Management at the University of Ottawa. She holds a University Research Chair in Gender, Diversity and the Professions. She leads the Canadian Health Workforce Network and the Empowering Women Leaders in Health initiative. Dr Bourgeault has garnered an international reputation for her research on the health workforce, particularly from a gender lens. Past projects have examined the migration and integration of health workers from a comparative perspective and on primary and maternity care workforce issues. Recent projects focus on care relationships in home and long-term care, and on psychological health and safety of professional workers.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**