

Changing primary care performance measurement by getting started

Thesis submitted in accordance with the requirements of the University of Liverpool for
the degree of Doctor in Business Administration by

Carol Anne Teresa Mulder

June 25, 2018

Acknowledgements

Thanks to all the members of the Association of Family Health Teams of Ontario and especially the Quality Improvement Decision Support specialists for joining me (or rather allowing me to join you!) on this journey. This is action research. It completely depends on action. I am so grateful for your courage, wisdom and vision in taking whatever action you could, whenever you could. Even more, I am grateful for the gift of your trust in sharing your story with me as we learned together how to get just a little bit better with each step along the way.

I am also grateful to those sharing their expertise in making D2D mean something: Ross Kirkconnell for steering everything from the start; Rick Glazier for knowing so much about measurement and being so keen on D2D anyway!; Jen Rayner for being such an excellent model of sharing and leading from the ground-up; AFHTO's Board (especially Randy Belair, Marg Alfieri, Rob Annis) for your enthusiastic and unwavering support of the hard work of measuring to improve, and Angie Heydon for making the space for me to get started with all of this.

Thanks to my supervisor Allan Macpherson. Maybe it was not your first time *supervising* doctoral research in a yurt but it was my first time *doing* it so I appreciated the company from afar.

Thanks to the assembled crowd of unusual suspects who added wit and/or wisdom to the story: Marc Croteau for keeping the figure fat; Donna Brown for helping me keep the figure thin, Liz Head for being there before the beginning and sticking around to the end anyway, Kathleen Delgatty for faith that this too would pass, Brenda Bonnett for introducing me to my

first 2X2 table and Jim Avon for, once again, fuelling my academic efforts with fantastic homemade curry.

Thanks finally to my partner Paul Smith, who kept the teapot warm and everything else right side up. As my mother used to say: Elke pot heeft een deksel. Thanks for being there. Those of you who came to understand the composite quality indicator through the quintessentially Canadian concept of the hockey pool (an analogy that sadly did not make the final cut in the thesis!) also have Paul to thank!

And now, with 50,000 of my words ahead, the last word here goes to the poet ee cummings:

i thank You God for most this amazing day:
for the leaping greenly spirits of trees and a
blue true dream of sky; and for everything
which is natural which is infinite which is yes
(i who have died am alive again today, and
this is the sun's birthday; this is the birth
day of life and of love and wings: and of the
gay great happening illimitably earth) how
should tasting touching hearing seeing
breathing any-lifted from the no of all
nothing-human merely being doubt
unimaginable You?
(now the ears of my ears awake and
now the eyes of my eyes are opened)

Abstract

This action research study was conducted as part of an ongoing performance measurement initiative in an association of primary care teams in Ontario, Canada. The problem addressed was the challenge of increasing participation in performance measurement. The research question addressed was: What happens when a novel approach to measuring quality/ demonstrating value is introduced in my organization? My view in this action research considered change as a continuous phenomenon. This is consistent with my operational mandate as a scholar-practitioner and my own orientation as a reluctant constructionist. The literature informed the development of a framework to guide the analysis of data in this study.

This action research was based on a developmental evaluation using qualitative tools for data capture and analysis. It ensured a critical perspective by orienting around the reflective questions: “What, So What, Now What”. The study unfolded in a cyclical way starting with the launch of the artifact, the experience with the first iterations, reflection on the experience to generate recommendations for action, the experience with implementing the actions (or not, as events unfolded) and finally reflections and general considerations for next steps with the initiative. The data sources included: results of surveys, minutes and materials for governance committees, email conversations between staff, members and stakeholders of the organization, my own observations and performance reports generated by the ongoing measurement initiative. Data were analysed using template analysis.

The artifact in this study was Data to Decisions (D2D) a multifaceted initiative that involves member engagement, supporting materials, a performance report and communication. The study showed that participation in D2D was high and that the initiative

was generally considered to be successful. The key themes emerging from the experience with D2D were: a focus on relationships; a dynamic of help-seeking and self-reliant behaviour, a range of perceptions of priority and a clear intent and ability to “get started” with measurement. Reflections on these themes generated actions, the fate of which was described and reflected on in the final phases of the action research study. Implications of the data were presented for consideration by the organization as the ongoing measurement work continues, independent of this action research study. In this way, the study contributed to the organization’s ability to support ongoing measurement and improvement of performance. The observation that relationships are more important in participation in performance measurement than the actual indicators being measured is a useful contribution to professional knowledge regarding performance measurement in primary care.

The thesis concludes with a reflexive moment in which I described the purpose and nature of reflexivity involved in my action research and summarize my role as a scholar-practitioner and my reflections on the project as a whole.

Contents

Acknowledgements.....	2
Abstract.....	4
Table of figures.....	10
Table of tables.....	12
Glossary of Terms.....	13
Chapter 1 Introduction.....	15
Context for the organization	16
Background of measurement of quality in primary care.....	18
The Data to Decisions initiative.....	22
The problem being investigated.....	24
Problem statement.....	27
Research question	27
Structure of thesis	28
Chapter 2 Literature review	30
Introduction.....	30
Routines	30
A view of Change as Episodic and Planned	34
Planned change: implications of, and for, distributions of power in organizations.....	35
A Continuous View of Change	37
Change agency and resistance	39
Conversation.....	40
Crisis.....	41
Disruption.....	43
Action and Learning.....	45
Power.....	46
Summary of a continuous view of change	49
Conclusions.....	49
Implications of Literature Review on Research Design	52
Chapter 3 Methodology	57
General approach	57

Setting	60
Phases of the study	61
Methods.....	64
“What?”: Data collection	64
Data sources	65
“So What?” -- Reflection and Learning	74
“Now What?” -- Preparation for the next actions	74
Ethics approval.....	75
Chapter 4 Phase 1 of Action research project – the artifact	77
Preamble	77
Member engagement.....	79
Governance.....	79
Indicator selection process:	79
Participation	81
Supporting materials for D2D	81
Content of report	83
Team context.....	88
Communication	91
Summary of Phase 1: The artifact.....	96
Chapter 5 Response to first 3 iterations of D2D (Phase 2 of action research study).....	97
Preamble	97
Process and structure.....	99
Participation	100
Structures.....	102
Physician Champions	102
QIDS Specialists	103
Team characteristics	103
Performance on primary care quality measures	107
EMR maturity.....	107
Individual Indicators	108
Quality composite measure	109

Cost.....	111
Quality and Cost.....	112
Performance measures summary	113
Conversations.....	117
Overview of data extraction process	117
Aggregate dimensions/master themes	118
Summary of conversations.....	126
Phase 2 summary	131
Chapter 6 Phase 3: Preparation for 4th iteration.....	132
Preamble	132
Reflections and implications for next steps: Relationships	133
Reflections and implications for next steps: help-seeking and self-reliance	136
Reflections and implications: diversity of perceived priority of D2D	139
Reflections and implications: “getting started” in QI activity	142
Summary	146
Chapter 7 Phase 4: Response to 4th iteration of D2D (final learning phase)	150
Preamble	150
Changes to D2D based on initial experience	151
Relationships	152
Help-seeking and self-reliance	153
Diversity of perceived priority of D2D	154
Getting started at QI	155
Other changes for D2D 4.0	157
Participation and performance in D2D 4.0	161
Conversations.....	162
Experience with actions aimed at addressing Relationships	162
Experience with actions aimed at addressing Help-seeking and self-reliance	164
Experience with actions aimed at addressing Perceived priority of D2D	164
Experience with actions aimed at addressing Getting started	166
Summary of Phase 5	167
Chapter 8 Phase 5: Reflections on the 4th iteration of D2D (final learning phase)	170

Preamble	170
Reflections	171
Relationships	171
Help seeking and self-reliance	174
Perceived priority of D2D	176
Getting started	177
Novel measure of quality	182
Summary	182
Chapter 9 Reflexivity in this action research project.....	187
Preamble	187
Purpose of reflexivity.....	187
Data sources for reflexivity	189
My role as scholar and practitioner.....	190
Reflections on my reflections	191
Apples and penguins: action learning and research	191
The fractal nature of D2D	193
My story is, apparently, my story.....	195
Summary of reflections on my reflections	196
Appendix 1. Data dictionary	198
Appendix 2. Coding template for action phase 1, Chapter 5	199
Appendix 3. Coding template for action phase 2 (experience with D2D 4.0), Chapter 7	202
Appendix 4. Abstracts submitted and accepted for presentation at various conferences	203
References.....	207

Table of figures

Figure 1-1: Timeline of D2D events	23
Figure 2-1:Knowledge, Action, Behaviour model of change (Source: Prager, 2012)	35
Figure 2-2: Queen Street bridge, Toronto, Ontario, Canada (Mulder, 2015)	37
Figure 2-3: Literature-based research framework to guide data extraction for this action research project	53
Figure 3-1: Situation of action research cycles in the context of operational activity of my organization and academic activity of my thesis	63
Figure 4-1: Summary of D2D indicator selection process.....	80
Figure 4-2: Evolution of the process of generating the composite quality score.....	86
Figure 4-3: Evolution of the EMR data quality composite measure.....	90
Figure 4-4: Readership of eBulletins	93
Figure 4-5: Layout of first iteration of D2D report.....	94
Figure 5-1: Contribution to D2D by iteration	101
Figure 5-2: Primary Care Practice Report uptake after introduction of team-level access.....	102
Figure 5-3: Team Climate Inventory scores	104
Figure 5-4: Proportion of teams with frequent (monthly or more often) conversation with physicians about performance.....	105
Figure 5-5: Percent of teams using standard tools and definitions to access data for selected D2D indicators	105
Figure 5-6: Change in performance over iterations	109
Figure 5-7: Homemade cupcakes provided by AFHTO staff member to celebrate launch of D2D 2.0.....	125

Figure 7-1: Gift of staff of one team to leader to celebrate achievements within the team

(i.e. local celebration) 157

Table of tables

Table 3-1: Phases of the action research study	61
Table 3-2: Timeline of D2D events	66
Table 3-3: Data sources used to reflect stakeholder perspectives.....	69
Table 3-4: Example of coding template.....	71
Table 4-1: Core performance indicators in D2D iterations.....	84
Table 5-1: Guidance for review of Chapter 5	99
Table 5-2: Regression of quality on per capital health care cost (Mulder et al., 2016)	113
Table 5-3: Summary of observations, reflections and implications of process and structure measures.....	115
Table 5-4: Summary of observations, reflections and Implications of conversations	128
Table 6-1: Themes and actions emerging from the review of data in Phase 3	149
Table 7-1: Fate of recommended actions for D2D 4.0	160
Table 7-2: Dimensions emerging from review of data in Phase 4.....	168
Table 8-1: Reflections and actions from earlier iterations of D2D, reflections on D2D 4.0 and implications for consideration by organization for future iterations.....	184

Glossary of Terms

AFHTO: Association of Family Health Teams of Ontario, the advocate, network and resource for team-based primary care in Ontario.

D2D: Data to Decisions, a voluntary performance measurement report developed and produced by members of the Association of Family Health Teams of Ontario

ED: Executive Director of a Family Health Team, one of a leadership triad that includes the chair of the Board of the Family Health Team and the medical lead, a physician.

EMR: Electronic Medical Record owned and used by teams to maintain patient health records

FHO: Family Health Organization, a physician-remuneration model common among Family Health Teams

FHT: Family Health Team, a formal model of team-based primary care receiving government funding for administrative and interdisciplinary healthcare professional staff

HQO: Health Quality Ontario, the provincial advisor on quality in health care.

IHP: Interdisciplinary healthcare professional including professions such as dietitians, social workers, occupational therapists, pharmacists etc.

MOHLTC: Ministry of Health and Long Term Care of the government of Ontario

PCR: Primary care report, a summary of performance based on administrative healthcare data, managed by Health Quality Ontario

QI: Quality Improvement

QIDS: Quality Improvement and Decision Support, the name of the program responsible for performance measurement within the Association of Family Health Teams of Ontario

QIDS specialist: Staff member of Family Health Team responsible for providing Quality Improvement and Decision Support, with the help of the QIDS program of Association of Family Health Teams of Ontario

Chapter 1 Introduction

“The journey of a thousand miles begins with a single step” (Laozi, n.d.)

Primary care has been identified as the foundation of a sustainable healthcare system (Starfield, 2009). As part of its quest to protect Ontario’s publicly funded healthcare system now and in the future, Ontario has invested considerably in team-based primary care (Aggarwal, 2009). Included in this investment is an increasing interest in measuring performance to support improvements in efficiency and quality of care. This, however, is a challenge. Finding meaningful ways to practically measure the breadth of primary care has proven difficult, even though there is a well-developed history of performance measurement and quality improvement in healthcare.

This research explores what happens when primary care teams act to address this challenge. It focusses on the first step on the journey towards manageable, meaningful measurement of primary care performance. While approaches to performance measurement are of interest in this story, they are not the primary focus. Instead, the object of attention is how getting started with performance measurement can change the journey, the destination and those who travel together along the way.

This chapter outlines the context for this research. First it describes the healthcare sector and the organizational setting for the research and my role within that. It also summarizes the background of the issue being investigated: measurement of quality in primary care. It presents the problem statement and research question. Finally, it lays out the structure of the subsequent sections of this thesis.

Context for the organization

My role in primary care is to lead the Quality Improvement and Decision Support (QIDS) program of the Association of Family Health Teams of Ontario (AFHTO). Family Health Teams (FHTs) were introduced in 2005 as part of a government transformation agenda (Conference Board of Canada, 2014). Teams are distinct from the historical approach to primary care, which was largely delivered by individual family doctors, sometimes with the help of an office nurse. Teams usually include multiple physicians and other healthcare professionals such as social workers, occupational therapists, and pharmacists, to name a few. The Conference Board of Canada (2014) noted that the FHT initiative was a response to the increasing awareness of the downstream benefits of better quality and cost-effectiveness associated with a strong primary care system. Inter-professional team-based care was noted as a key attribute of strong primary care system. Team-based care was proposed as an effective and efficient foundation for a sustainable publicly-funded healthcare system in Ontario. Between 2005 and 2015, 184 FHTs teams were set up, covering approximately 25% of Ontario's population¹. AFHTO was born out of a desire of the teams to have a collective voice, primarily to advocate on their behalf and help them demonstrate their value in terms of better quality and lower system cost. All interdisciplinary team-based primary care organizations in Ontario are eligible, but not required, to join AFHTO. AFHTO represents all but 4 or 5 FHTs across the province (depending on the year).

One of the key strategic priorities outlined in AFHTO's 2013-15 strategic plan was to measure and improve the quality of care provided by its members as part of its overall mission to

¹ In addition, there are two other models of team-based primary care in Ontario (i.e. Community Health Centres and Nurse-Practitioner Led Clinics). The nearly two hundred teams set up according to these two models cover roughly 25% of the Ontario population.

advocate and support members in improving and delivering optimal inter-professional care (AFHTO, 2013). On the strength of this mission, AFHTO could negotiate funding for a quality improvement program from the ministry. The funding for the Quality Improvement and Decision Support (QIDS) program was intended by the Ministry of Health and Long-Term Care (government) to facilitate compliance of AFHTO members with annual reporting of a Quality Improvement Plans. From AFHTO's perspective, the purpose of the program was broader. It was intended to support progress on AFHTO's strategic priority to advance measurement and demonstrate the value of primary care teams.

The majority of the QIDS program funding covers the salaries of QIDS specialists hired directly by AFHTO members. Each QIDS specialist supports a partnership of several AFHTO members. A total of 35 full-time positions were approved among 150 of the 184 AFHTO members. The remainder of the QIDS program funding covers 3.5 full-time staff in the central AFHTO secretariat, of which I am the lead. The decentralized deployment of the QIDS program reflects AFHTO's recognition of the autonomy from its members. Participation in any initiatives launched by AFHTO or the QIDS program is therefore voluntary, as is virtually every other aspect of the association, including membership, as noted above. Consequently, AFHTO staff are very committed to meeting expectations and ensuring the ongoing interest of members in the organization.

Despite the investment in FHTs over 10 years ago, the perceived concerns about quality and cost of care persist. According to measures commonly valued and quoted by healthcare system decision-makers, the quality of primary care in Canada and Ontario lags other nations (CIHI, 2016; HQO, 2014). Specifically, Ontario performs poorly on access to care and aspects of primary care infrastructure, measures that are considered relevant to the

government of Ontario, if not to providers and patients. There is also a narrative among decision-makers that, while team-based care is still a good idea for achieving good quality primary care, teams, as they currently are configured (i.e. FHTs), are too expensive (Grant, 2015). The government's most recent effort to transform the primary care sector, the so-called "Patients First" Bill 41 (Ontario, 2016), seeks to expand access of Ontarians to team-based primary care without expanding teams. This signals the government's espoused belief in expanding, measuring and reaping the benefits of team-based primary care for all Ontarians.

In summary, there is a long-standing drive to improve quality and efficiency of primary care in Ontario. There is also a continuing commitment to interdisciplinary team-based care to achieve that. In fact, one of AFHTO's strategic priorities is to demonstrate the value of team-based care through measurement and improvement of quality. This priority formed the basis of the intervention studied in this research.

Background of measurement of quality in primary care

The phrase "You can't improve what you can't measure" yields over 4.5 million hits in a Google search. The important role of measurement in improvement is at the heart of the work of Drucker (1973) and Deming (as described by Howell, 2006) who are widely acknowledged as leading advocates of performance measurement to lead to quality improvement. Nevertheless, there do not yet appear to be adequate processes to measure quality in primary care in Ontario (Ashcroft, 2014; Hutchinson & Glazier, 2013), despite the longstanding and highly publicized interest in improving it. This is curious as there is an exponentially increasing amount of data available to support measurement and improvement in primary care (de Lusignan & van Weel, 2006). In Ontario, approximately 80% of primary providers have Electronic Medical Records (EMRs) as of 2015 (ehealthOntario, 2015) compared to 25% in

2003 (Keshavjee, 2007). There is also evidence that the quality of the data is sufficient to support measurement for improvement and research (Greiver et al., 2014). However, this has not yet facilitated more measurement or demonstrably improved care (Greiver et al., 2011).

There is considerable literature around best practices for measurement in healthcare. Leadership, particularly physician leadership, is widely acknowledged as a prerequisite for measurement in healthcare (Brown, 2010; Crabtree et al., 2005; Gallagher et al., 2010; Kirchner et al., 2012; Lee et al., 2014; Marsteller et al., 2011; Wolfson et al., 2009). AFHTO acknowledges this, with 5 physicians among 13 Board members of the association and physicians frequently in the majority on boards of local primary care teams. The provincial government is recruiting physician leaders for each of more than 70 healthcare regions across the province.

Another enabler of measurement is the establishment of standardized measures (HQO, 2014). These measures must also be actionable (Ivers et al., 2014). The provincial government responded to the need for data by investing in a Primary Care Performance Measurement framework (HQO, 2014) with emphasis on alignment with international standards and development of data collection and reporting infrastructure. Consultation began in 2012. Over the course of 2.5 years, this process produced a library of over 200 indicators, 70% of which were not currently implementable (HQO, 2014, p 13). The UK (Guthrie, 2008) and USA (Zaslavsky et al., 2002) have developed similar voluminous libraries of indicators. Ironically, Quality Improvement (QI) leaders are now complaining that there are too many indicators (Cassell et al., 2014; Stempniak, 2013; Zaslavsky et al., 2002). The perception has evolved to suggest that it is now the plethora (not the absence) of

data and indicators that is impeding the sector’s ability to advance measurement and improvement.

Access to good quality data is also identified as a key enabler for measurement. Part of Ontario’s response to this need was the production and release of Primary Care Reports available free of charge to all of Ontario’s approximately 10,000 family physicians. Ontario has also successfully deployed primary care EMRs for over 80% of primary care physicians (eHealthOntario, 2015). EMR data are more timely, more patient-centered, and more provider-specific than administrative data sources, and thus are potentially more useful to providers. Researchers have demonstrated that the quality of the EMR data can be measured and is sufficient for research, measurement and improvement purposes (Greiver et al., 2014).

Rewards such as “pay for performance” schemes are considered to be helpful in increasing participation in efforts to measure and improve healthcare outcomes (Lindenauer et al., 2007). A culture that makes it easy for providers to identify areas for improvement without being blamed or shamed as “poor performers” has increasingly been identified as being instrumental in facilitating measurement and improvement (Ivers et al., 2014). Ontario has implemented bonuses for achieving targets in EMR implementation in some primary care processes, such as cancer screening. Because there are no penalties for failing to achieve specific targets, and no public release of performance data, Ontario could arguably be considered to have a blame-free culture regarding primary care performance.

One challenge with measuring and improving performance is the nascent state of the definition of quality in primary care. Consider just one aspect of primary care: access to care. The focus of international and provincial decision-makers on patients’ ability to get an appointment on the same or next day when they are sick (HQO, 2014) has been shown to

have a perversely negative impact on patient experience (Kiran & Obrien, 2015). The governmental focus on Emergency department use as a measure of access to primary care is out of line with recent reports that use of Emergency department is almost completely independent of availability of primary care services (Green et al., 2016). Patients' views of access to primary care are quite different from either of these (Patients Canada, 2015; Family Medicine for America's Health, n.d.). Given the multiple ways of understanding even this one aspect of primary care (that is, access), it is not surprising that there is considerable diversity of opinion around the selection of indicators to reflect overall quality in primary care.

Given the debate regarding which individual indicator(s) best represent quality, composite measures represent a potential solution. Because they can incorporate many measures, they have the potential to neutralize debate about which one(s) should be monitored. They can also reflect performance on many different indicators at the same time without creating the concomitant burden of tracking many indicators. For example, the Quality Outcome Framework used in the UK incorporates 150 indicators into a single score (Guthrie, 2008; Lester & Campbell, 2010). TRANSFORMATION, a Canadian initiative, is exploring the development of "Primary Health Care Performance Portraits" (McGrail et al., 2015), which incorporate patient experience and other traditional performance indicators into composite measures, but at time of writing had not yet progressed to the point of defining the components and calculation process for generating these composite measures. Another composite quality indicator is the Summary Quality InDex (SQUID), which is based on 36 indicators (Nietert et al., 2007). Composite measures are not without their critics. The main arguments against composite indicators appear to relate to the perceived difficulty in

generating them and the lack of interest and/or ability of providers to act on them, partly due to poor face validity (Kaplan et al., 2009; Scholle et al., 2008; Reeves et al., 2007).

Nevertheless, composite indicators remain attractive from a theoretical perspective because of their ability to represent a breadth of information in a single score or index, thus addressing the issue of proliferation of measures.

In summary, measurement of quality in primary care in Ontario remains at a rudimentary stage, with continuing debate about how to define quality and, from there, debate about how to increase participation in measuring it. Measurement, or lack thereof, therefore remains a rate limiting step for improvement in quality of primary care in Ontario. This research examines the problem of performance measurement, focussing on what happens when primary care teams get started with it specifically through the Data to Decisions (D2D) initiative, described in the next section.

The Data to Decisions initiative

One of the key initiatives of the QIDS program was the measurement initiative called Data to Decisions (D2D). D2D is described by the organization as a summary of performance of primary care teams on a small number of indicators identified by members as both meaningful and possible to measure. Figure 1-1 shows the timeline of D2D events from first to the sixth iteration (D2D 1.0 to D2D 5.0). The performance report is only part of the overall D2D initiative. The D2D initiative is composed of member engagement activities such as surveys, direct conversations and webinars, supporting materials to facilitate access to and use of data, the performance report itself, and multiple modes of communication to multiple audiences. This action research project considers the experience with the entire D2D

initiative, not just the performance report, even though that remains the mostly visible aspect of the initiative internally and externally.

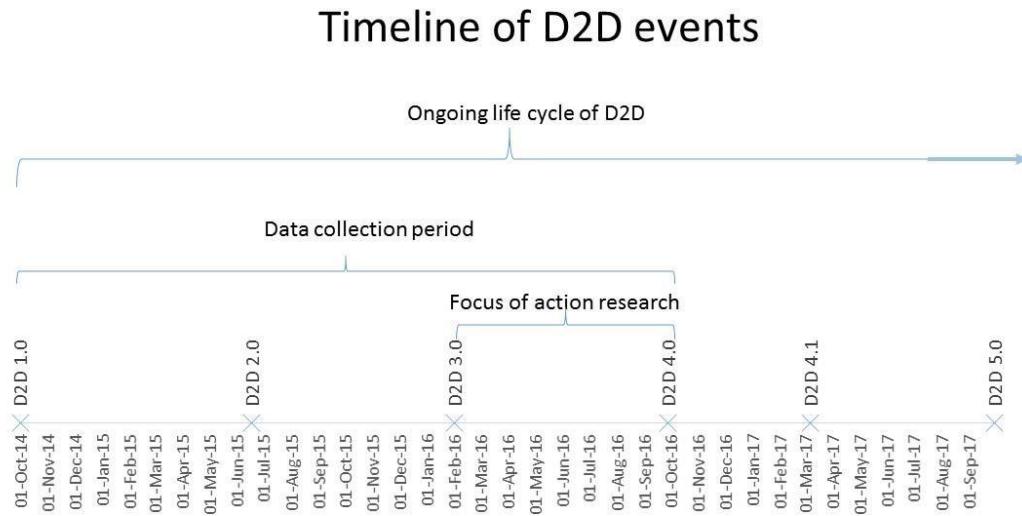


Figure 1-1: Timeline of D2D events

The action research project on which this thesis is based is superimposed on the pre-existing and ongoing action learning exercise of the D2D initiative. The action *research* study is distinguished from the ongoing action *learning* by the structured and rigorous methods used to learn from and guide action. Another distinguishing feature of the action research project is the intent to generate and publish knowledge emerging from the collective experience (Cassell & Symon, 2004). The action research study focusses on the interval between the first three iterations (D2D 1.0-3.0) which were complete prior to the start of the action research and the fourth iteration (D2D 4.0). The action aspect of this action research project centers on how the learnings from the first three iterations were incorporated into the deployment of D2D 4.0.

Certainly, the learnings from D2D 4.0 were also incorporated into subsequent iterations D2D 4.1 and D2D 5.0, released in March and September 2017, respectively. However, since these two iterations were released after the data collection period, the experiences related to them are not addressed in this study.

The problem being investigated

The problem with the lack of a credible performance measurement process is that it is threatening the viability of primary care teams. Without data demonstrating the value of interdisciplinary primary care teams, they are at risk of decisions that could restrict their funding and otherwise limit their ability to operate effectively.

The lack of evidence of the value of primary care teams cripples the sector's ability to negotiate effectively for human resources. Because of the perception that teams are too expensive (Grant, 2015), the government announced that it was pausing its investment in interdisciplinary care teams. This prevents recruitment of additional physicians to the teams. There is also a financial challenge in recruiting interdisciplinary healthcare professionals (IHPs) to teams because of poor salary equity with other healthcare settings (AFHTO, 2012). Lack of data demonstrating the value of the service provided by teams makes it hard to argue effectively for more funding for human resources, threatening the very existence of the teams. As outlined above, there are potential solutions to the problem of lack of credible performance measurement in primary care. Ontario's healthcare system has clearly responded to the evidence regarding enablers for quality improvement. Based on this, Ontario could be expected to be making good progress in measurement and improvement. Yet this is not the case. The reasons for that are not obvious.

One possible explanation for the lack of impact of enablers on quality is that participation in them is still low at the front line. For example, although Primary Care Practice Reports were made available to all 10,000 family physicians, only about 600 had enrolled by the third year. An intervention with less than 10% penetration cannot be expected to achieve its goals.

Even with good participation in quality improvement activities, quality might not improve. Compliance with mammography guidelines failed to improve breast cancer outcomes (Hall, 2014) and rigorous attention to blood glucose monitoring does not always improve outcomes of patients with diabetes (Qaseem et al., 2014). These observations raised questions about how valid it is to assume that changes in the level and nature of participation in QI activities (if any) could result in improved outcomes.

Another contributing factor to low impact and participation is the lack of requirement (either through positive or negative reward) to demonstrate value for money invested in primary care. The literature is mixed about the impact of financial incentives on performance, with clear signals of positive impact on EMR data standardization behaviour (CMA, 2014) and little or even negative impact on quality in response pay-for-performance plans (De Silva & Bamber, 2014; Hutchison & Glazier, 2013). For these or other reasons, there are no formal expectations for primary care providers to report on specific clinical indicators, or achieve minimum levels of performance, in any of the reports that primary care teams are required to submit.

The perceived relevance (or lack thereof) of the data being made available to providers may be another factor contributing to lack of momentum in measurement and improvement. Wide-spread and easy access to data that are not meaningful to providers (e.g. same/next day

appointment availability or Emergency department use indicators described above) is not likely to lead to improvement.

Finally, an important consideration in the lack of participation in improvement activity is the fundamental belief about the extent of the problem. The field of continuing professional development in medicine has long struggled with the paradox that those most in need of building their knowledge and skills are least likely to seek support in doing so (Davis et al., 2006). The same paradox applies here: participation in measurement is necessary to demonstrate the need for further engagement in measurement.

The difficulty Ontario is facing with respect to measuring and quality in primary care, therefore, is not a lack of potential solutions. Nor is it failure to implement, or at least attempt to implement, some of the solutions, as noted in the preceding examples. Instead, all the above suggests that, while the literature may be correct in identifying specific characteristics as enablers of measurement and improvement, they are not guarantees of the desired outcome. The problem may not be the various processes or tools or approaches to performance measurement, which could truly be excellent in themselves. Instead, the problem may be more about supporting the organizational change inherent in starting, and responding to measuring performance.

The concept of “best” practice assumes that it is possible to define a right or wrong approach that is applicable to all situations. Constructivist theory is based on an alternative assumption that there is no one truth or one “best” approach. Instead, everyone makes their own meaning of what is “best” for them (Aguinaldo, 2004). Truth, in the conventional sense, is less meaningful than the extent to which people are persuaded by information (Patton, 1999). The incomplete penetration and effectiveness of what are widely held to be best

practices embodies what Churchman (1967) referred to as a “wicked problem”. There is no one way to fix it, only better or worse alternatives and the solution chosen can generate other problems (Grint, 2005). As such, it implies that additional strategies to improve engagement with measurement need to be considered. The lens of constructivist theory suggests that equipping people to find ways to act that make sense to them could be an alternative or additional strategy, rather than continuing to try defining and implementing the “best” solution for everyone. To that end, the focus of this effort to advance measurement is not so much on developing the measurement tools and activities themselves, but on the process of change related to introducing and encouraging measurement in organizations.

Problem statement

The problem I am focussing on is how to increase participation of primary care providers in measuring and improving primary care. The problem persists in the face of “best practices” described in the literature, some of which are already deployed to greater or lesser extent within the organization.

Research question

The intent of this research is to increase ownership of and participation in the measurement problem by facilitating reflection and learning through the introduction of a novel way of measuring and demonstrating value in primary care.

The primary research question addressed is: What happens when a novel approach to measuring quality/demonstrating value is introduced in my organization? Given that the organization’s experience with measurement preceded and will continue independent of this research, a secondary research question is: What does it take to continually increase participation in measuring and improving quality of primary care?

Structure of thesis

This thesis is organized into several sections as described below. References and appendices are included in separate sections.

Literature review: The literature review summarizes how existing theory of change contributes to our understanding of how an organization experiences a change. It focuses on how the experience can generate useful learning to guide subsequent interventions to support measurement and improvement.

Methodology: The methodology section describes the framework guiding the analysis of data. It outlines the data sources and the approach to analysis of the qualitative data, based largely on template analysis (King, 2004). The quantitative data considered in this research came in the form of summary reports in archival documents, not the actual quantitative data themselves. For that reason, the methodology is focussed on qualitative techniques.

Findings: The findings part of this thesis includes observations regarding the actions taken related to the artifact (that is, D2D) as well as reflections on those actions. The findings begin with a description of the artifact. Because there are two action phases (the first 3 iterations of D2D considered together and the fourth iteration), there are also two reflection phases. The nature of data differs slightly between phases and includes diverse sources such as survey results, performance data, archival documents describing the D2D initiative as well as conversations describing the experience of AFHTO's members and stakeholders with D2D.

Reflections: The reflections associated with each of activities involved in the ongoing evolution of D2D are presented along with the data on which they are based. In addition,

there is a separate series of reflections specifically focussed on my impact on the initiative and vice versa.

Chapter 2 Literature review

Introduction

As discussed in Chapter 1, the main problem being addressed in this thesis is how to support participation in performance measurement in primary care. It does not address the actual design of measurement tools in general, nor the tool being studied here. To that end, my review of the literature is focussed on issues related to change management, not specific characteristics of performance measurement strategies. My review of the change management literature focusses on change as a continuous phenomenon (Weick & Quinn, 1999) with a brief review of literature supporting a vision of change as an event, to provide context. This literature review will not make a case for the rightness or wrongness of either vision because that requires first making a case for the rightness or wrongness of a person's way of understanding knowledge. That is beyond the scope of this review. Instead, this review will illustrate these visions of change along with their implications regarding power distribution in organizations. The review starts with a brief description of organizational routines, since they are the target of the two different visions of organizational change. The review concludes with description of the theoretical grounding of this thesis research in a view of change as a continuous phenomenon. This review also presents a framework informed by various concepts related to the theory of continuous change that guides the action research study.

Routines

Orlikowski (1996) argues that organizational behaviour consists of regularized practices, or routines, that are constantly evolving to maintain equilibrium. By "routines", she is referring to observable patterns of behaviour like standard operating procedures as well as less tangible behaviours, such as the norms of intra and interpersonal interactions. She

suggests that these routines are a mechanism for storing the results of individual learning in organizations, ensuring that the knowledge gained by an individual is available to everyone in the organization via everyone's participation in the routine. Berends et al. (2003) observed that routines can persist in the absence of the individual involved in the original learning because the norms and behaviours are taught (explicitly or otherwise) to new members as they become part of the organization. Routines define and therefore effectively reinforce the status quo (Nelson & Winter (1982), Levitt & March (1988 in Edmondson et al., 2001)). For this reason, routines affect an organization's experience of change.

One such routine is the assumed role of rationality in decision-making. So-called rational decision-making fits Argyris' (1990 in Adams, 1994) definition of routine in that it involves inconsistent messages whose inconsistency is denied. Irrationality (or even the possibility of it) is undiscussable. Consequently, as Vince & Broussine (1996) observed, privileging the concept of rationality makes it acceptable to dismiss any uncomfortable ideas by simply labelling them as irrational. They note that this creates an easy avenue for dealing with distracting or emotional reactions, whether they reside in others or are internal to the decision-maker. Menzies-Lyth (1990 in Vince & Broussine, 1996) concluded that rationality could therefore make it easier for decision-makers to deal with their own uncomfortable anxiety. They found that this might be particularly welcome to managers fixated on task completion and strategic problem-solving who invariably find emotions and feelings difficult to articulate. As Brown & Jones (2000, in Smith & Elliott, 2007) notes, rationality also creates more space for delusion in the interests of avoiding embarrassment. Palmer & Dunford (2008) concluded that delusion is inherent in rationality because you can't erase a person's lived experience with all the emotions that entails, by edict, even if you wanted to.

The delusional capacity of rationality can help decision-makers avoid the discomfort of change in a way that helps them avoid any embarrassment related to their decision. Since avoiding embarrassment has been identified by Argyris (1996) as a primary driver of organizational behaviour, the high affinity for rationality in organizations is not surprising. In this way, respect for rationality serves as an example of a norm or routine that discourages change by discouraging the uncomfortable questions that expose assumptions behind existing behaviour and decisions.

Another routine that potentially plays a role in organizational change is organizational silence. Morrison & Milliken (2000) define organizational silence as a consistent, subconscious rule that governs what is and is not spoken of in the organization. Organizational silence has been observed to contribute to unity, agreement and consensus, which are accorded a positive valence in many organizations (Wagner, Pfeffer, & O'Reilly, 1984). In contrast, disagreement is considered to be impolite (Fook & Askeland, 2007). Morrison & Milliken (2000) observed that questioning or articulating assumptions of leaders is perceived as lack of trust and therefore not welcomed, or rewarded, by those with power in organizations. This norm of unity and silence (as opposed to disagreement) has been observed to be even stronger in the face of crises or threats (Staw et al., 1981). Brockner & James (2008) showed that even relatively small crises, such as the risk of embarrassment, prompt increased rigidity in routines and thus discourage divergent thinking. Janis (1973) described the phenomenon of groupthink, in which the silencing of dissent limits an organization's ability to deal effectively with crisis and change. Morrison & Milliken (2000) concluded that, like most routines, organizational silence serves a purpose (that is, unity) at the same time as it creates risk in the form of groupthink.

A third important type of routine in consideration of organizational change is the defensive routine. Defensive routines have been described by Argyris (1996) as behaviours used by individuals to defend against embarrassment and threats to control. A variety of defensive routines described by many different authors have been compiled by Vince & Broussine (1996). Some of the examples in their compilation include a tendency to default to reactions that have provided some level of comfort or security previously and over-compensating in the opposite emotional direction to the threat (e.g. go on the offensive when feeling attacked). Argyris (1999) contend that defensive routines are the norm in organizations and they contribute to our collective inability to see the difference between what we aspire to do (that is, our espoused theories) and what we actually do (that is, our theories in use). They concede that it is possible to break these routines, but argue that this requires examining deeply held values and assumptions in the face of a systems that reward blindness. They further observe that this takes more courage and skill than most people have. Consequently, defensive routines interfere with the ability of individuals and organizations to learn and change.

Rationality, organizational silence and defensive routines are thus three example of routines that are relevant when considering organizational change. Not all routines serve to encourage stability, however. Routines have also been shown to facilitate change. Feldman & Pentland (2003) and Deken (2016) describe a dynamic between ostensive and performative aspects of routines. They suggest that the historic basis of routines based on what has been learned and embedded in the organization (the ostensive aspect) also has a current or real-time component as the routine is actually conducted (the performative aspect). In the ongoing adaptation of humans to their environments, they argue that the performative aspect of routines can contribute to change. Feldman & Pentland (2003) also describe meta-routines

such as total quality management and continuous improvement which are patterns of behaviour intended specifically to support change. However, defensive routines are particularly relevant to this work, given their persistent and unexamined nature, and the extent to which they are deeply embedded in daily behaviour, which makes them very difficult to change. There is thus a tension between the stabilising and innovative aspects of routines that are connected to the challenges associated with change. Therefore, the theoretical treatment of the role and impact of routines is examined in both visions of organizational change presented below.

A view of Change as Episodic and Planned

As described by James (1909/1996 in Tsoukas & Chia, 2002), the Aristotelian and Platonic view of change from nearly 2500 years ago sees change as the exception, the unusual and even the less worthy state of an otherwise fixed environment. Stability from this perspective is fixity or absence of change. Change requires this fixity be disturbed or unfrozen and that this disturbance creates the opportunity to move. Burns (2004) reports that Kurt Lewin is credited with describing the process by which this type of change unfolds using a 3-stage model referred to as “unfreeze-move-refreeze”, opposite to the model of continuous change. Once the movement has occurred, the organization then resumes its new steady (that is, frozen) state. The literature on planned change is replete with what Collins (1998 in Randall, 2004, p. 145) referred to as “n-step models of change” (Arrata (2007), Bahamon et al. (2006), Kotter (1995), Myrvold (2011) and Wiest (2006)). They are characterized by: rationality, sequential approaches and a “generally upbeat and prescriptive tone”. (Collins, 1998 in Randall, 2004, p. 145). Randall (2004) suggests that the attraction of these models is the perception that they make the daunting task of change easier. The idea that the change

leader carries the risk for the success of the change may explain the vision of a heroic change leader that Palmer & Dunford (2008) observe is prominent in the literature about organizational change. One of the challenges faced by these heroes is resistance to change. Not surprisingly, resistance is framed as a bad thing (Palmer & Dunford, 2008) and a problem to be solved in accomplishing change. (Vince & Broussine, 1996). Another characteristic of the vision of change as a planned event is the timing of behaviour change which is at the end of the unfreeze-move-freeze cycle. An example is the “persuasion theory” of change described by Prager (2012), commonly referred to as the “Knowledge, Attitudes, Behaviour” model in the popular literature. In this model, uptake of the behaviour is the last step in the change process (see Figure 2-1).

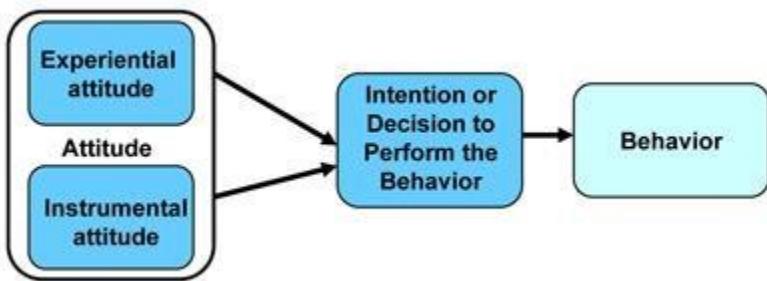


Figure 2-1:Knowledge, Action, Behaviour model of change (Source: Prager, 2012)

In summary, a view of change as a planned episode frames change as the exception from the normal stable state of an organization. Such a world view positions change as something that needs to be carefully led to ensure successful return to stability. It follows, therefore, that the distribution of power in an organization is highly relevant to how planned change is perceived to take place.

Planned change: implications of, and for, distributions of power in organizations

One feature of planned change that needs to be considered in the context of power is the role of resistance. Van de Ven & Poole (1995) suggest that questioning change activities is

often labelled as resistance and therefore discouraged. This interferes with the deeper level of learning required to achieve second-order change. The extent to which those in power in the organization perceive questions as resistance therefore affects the depth of learning and nature of change the organization can achieve.

Another aspect of planned change that needs to be considered in the context of those who hold power in the organization is its vulnerability to inertia and resultant slow or incomplete implementation. Carmeli & Schaubroeck (2008) suggest that managers might be reluctant to alert organizations to problems for fear of being perceived as less effective leaders. A slow or delayed start to change may serve such managers by giving them time to distance themselves from the problem. The pace of change in an organization could be affected by the extent to which powerful managers in an organization are among those who prefer extra time and space to work out their position on the change Turnbull (2001).

The locus of change management is another aspect of change that needs to be considered in the context of power in organizations. Vince & Broussine (1996) argue that the planned approach to change epitomizes the desire for power over uncertainty. The paradoxical problem noted by Gioia & Chittipeddi, (1991) is that not only does the planned/controlled approach not work to eliminate uncertainty, it also makes it harder to see a compelling reason to change. Pearson & Clair (1998) observed that the proactive, planned approach and perception of control contributes even further to the strength of the comfortable routines bred by success and fuels the equally comfortable complacency regarding the need for change. Therefore, the comfort level of powerful decision-makers with uncertainty and their capacity to cope with the “psychological pain” (Kilduff & Dougherty (2000), Weick & Quinn (1999) and Armenakis & Bedeian (1999)) of change affects how change unfolds in the organization.

In summary, many of the features of planned change appear under the control of, or at least responsive to, those holding power in the organizations. The extent to which those holding power in the organization can cope with the discomfort of change therefore limits the extent to which the organization will embrace the deep transformation of second-order change. Unfortunately, the success achieved by those in power creates complacency about the need for change, thus reducing their interest in the pain of transformational change. Change as a planned event therefore depends on, and can be obstructed by, the interests of those who hold power in the organization. Nonetheless, viewing change as a planned event is only one way of considering how change happens. An overview of change as a continuous phenomenon follows.

A Continuous View of Change



Figure 2-2: Queen Street bridge, Toronto, Ontario, Canada (Mulder, 2015)

"This river that I step in is not the river I stand in". These words (originally uttered by Heraclitus, in 500 BC and now, ironically, carved into steel and concrete – see Figure 2-2) describe an alternate understanding of change as the natural state of organizations (Tsoukas &

Chia, 2002). Many authors (such as Lewin (whose body of work was summarized by Burnes, 2004) and Orlowski (1996)) argue that organizational stability is a phenomenon of equilibrium between ongoing responses of all members to the continuous pressures they are experiencing in daily life. In their analysis of Giddens' (1984) structuration theory, Berends et al. (2003) contend that the social environment of the organization (e.g. power, norms) affects how individuals learn and change. They describe behaviour in an organization as the sum of the usual practices of its members. Further, these authors note that change happens when aspects of the social environment *or* individual practices *or* routines change. The organization can therefore be considered to be both stable and changing all at the same time. Weick & Quinn (1999) describe a 3-step model for change freeze-rebalance-unfreeze. In this model, they propose that organizational change starts with a pause that is long enough to examine an organizational practice. This is followed by an intervention of some kind that rebalances factors affecting behaviour (e.g. social environment, norms, power). The final stage is the resumption of activity after the pause, which releases the organization to return to the usual state of what Tsoukas & Chia (2002) refer to as “organizational becoming”. In this view of change, Weick & Quinn (1999) propose that there is no end state. They propose that interventions are meant to restore equilibrium, which has been disrupted by improvisations in coping mechanisms by members of the organization, or what Orlowski (1996) called “situated change”. Contrary to how it might appear through its embrace of emergent ideas, Van de Ven & Poole (1995) suggest that continuous change is not aimless. They reference biological evolution in which observed variations (that is, improvisations) might appear random, but the selection process for the apparently random variant best suited for a purpose is deliberate. Based on this, Palmer & Dunford (2008) argue that change is therefore about

supporting learning. The following sections outline themes that describe how change as a continuous phenomenon happens. They include change agency and resistance, conversation, crisis, disruption, action and learning and, finally, power in organizations.

Change agency and resistance

When considering leadership of change as continuous, Weick & Quinn (1999) contend that it is the process of improvisation or disruption, not a single person or group, that is the originator of change. The spread of the change idea (that is, the implementation of change) is considered by Palmer & Dunforth (2008) to be less dependent on how well a change agent can persuade others to participate in the change and more about facilitating ongoing adaptation and improvisation. Change agency therefore is described as being focussed on nurturing (Palmer & Dunford, 2008) and being attentive to emergent changes (Weick & Quinn, 1999). The focus on emerging, ongoing change has implications for understanding what is described as resistance to planned change.

Thomas et al. (2011) frame the behaviours traditionally labelled as resistance (e.g. challenges or modifications to proposed solutions) as signals of engagement in the ongoing process of adaptation, which is at the heart of continuous change. Resistance is described by Ford et al. (2008) to be part of the process of change, not a problem to be solved. They argue that challenges and complaints can help advance change by keeping conversation about the proposed change alive. Oreg & Sverdlik (2011) observed that sometimes these conversations can be characterized by ambivalence. These authors, along with Piderit (2000), suggest that ambivalence does not imply lack of support for the change, but rather could represent a balance of simultaneously strong enthusiasm for and equally strong antipathy to the change. Piderit (2000) suggests that recognizing and embracing the true nature ambivalence may be

more important than trying to sell the change to minimize the risk of inadvertently shifting the balance away from the desired behaviour. One mechanism for doing so includes a suggestion from Ford & Ford (1994) to identify and capitalize on what is attractive to people regarding the proposed change. Strong feelings and narrative against a change (historically viewed as “resistance”), or even ambivalent reactions to the change, can be useful clues regarding the nature of forces pushing back and forth on the equilibrium of the organization, and thus be important tools in achieving positive change. Conversation thus plays a role in reframing resistance. The next section describes other roles for conversation in continuous change.

Conversation

One way of understanding the reaction to change may be through attention to conversations and other interpersonal interactions. Dervisiotis (2002) observed that conversations are important not only to complete tasks but also to share emotions and beliefs. Weick (1988) held that it is through saying things that we begin to know what we think. He argued that this awareness then feeds the next actions we take, including the next things we talk about. In this way, conversations are valuable vehicles for understanding and shifting what people believe. Ford & Ford (1995) position change as a part of conversation, rather than the other way around, in that conversation creates new realities based on beliefs and preconceptions we might not even be aware of until we talk about them. Having developed a new awareness through conversation, we are then inclined to pay more attention to it in our conversation (Berquist, 1993 in Ford 1999). Conversations thus are not only the way to create change in narratives and beliefs but are also the product of that change, in that people will talk about the new ideas they are paying attention to. To that end, conversation has been proposed as an intervention in itself (Macpherson et al., 2006). This is also consistent with principles of

feedback and audit described by Ivers et al. (2014) and Yeo (2013), in which feedback on performance (i.e. through conversation) is expected to increase interest, dialogue and, in turn, engagement in improving care. While it is clear that conversations are instrumental to change, many observers (e.g. Ford et al. (2008) and Dervitsiotis (2002)) agree that not just any conversations would do; they need to be ‘conversations-for-action’ in the words of Dervitsiotis (2002, p. 1088) or “conversations for performance” in the words of Ford & Ford (1995, p. 549). Palmer & Dunford (2008) reported that such conversations require advanced listening skills tuned to hear and pay attention to new possibilities. According to the theory of social construction of knowledge and Weick's (1988) concept of enacted sense-making, knowledge will evolve through such conversations and the subsequent small actions (or even non-actions) and undertaken by members of the organization. The role of conversations as a response and enabler of change is one of the key concepts guiding this action research study (see research framework at end of Chapter). The next section describes crisis as an inescapable trigger for conversation and thus change.

Crisis

Crisis changes things. Smith (1997) observed that without waiting for people to be ready for it, crises reveal errors in our thinking and open up what is normally unthinkable and not discussable for everyone to see. While crises have been reported to create massive and possibly irrecoverable damage, Smith & Berg (1987) suggest they can also create otherwise unheard-of space for second order learning. Ford & Ford (1994) suggest that if you can survive a crisis, it can be a useful signal of the need to try another way of thinking. For that reason, Dooley, (1997), Lewin (1951, in Weick & Quinn, 1999), Nonaka (1988) and Smith (1995) were among many who looked to create crises, stirring up emotions and organizational

turbulence in aid of positive organizational change. In other words, their advice can be summed up as: “If it ain’t broke, BREAK IT!” (Kriegel, 1991). Weick’s (1988) theory of enacted sense-making holds that organizations act in response to disruption to make sense of it. Carmeli & Schaubroeck (2008) note that a crisis can facilitate change most effectively if it is able not only to disrupt operations to the extent that the survival of the organization is in question, but also challenge the most basic assumptions of the members of the organization. It is in making sense of the disruption that potentially encourages a process of questioning, learning and enacting new practices (Weick, 1988).

This suggests that creating crises deliberately may be a way of instigating change. The process of creating crises to support change has been described by various writers. In his description of crisis as a competitive tool in car manufacturing, Kim (1998) framed crises in terms of the two Chinese characters used to describe the concept: danger and opportunity. Argyris & Schon (1978 in Dooley, 1997) describe crisis as the deliberate act of looking for and paying attention to differences (rather than consensus) to expose the mental models being used. The overall intent is to “to break open the shell of complacency and self-righteousness” (Lewin 1951, p. 371 in Weick & Quinn, 1999). Intentional ambiguity is another strategy described by Gioia & Chittipeddi (1991) to encourage stakeholders in an organization to question the usual way of seeing things. Dooley (1997) identified other deliberate, thoughtful ways to create a crisis or what he called a “far-from-equilibrium condition” (p. 79), such as exploring (or even just clarifying) organizational boundaries, creatively identifying assumptions through non-verbal descriptions of the system, and embracing (rather than dismissing) statistical outliers. The next section describes other approaches to achieving the disruption inherent in crises.

Disruption

Crises are not the only tools for disruption. In the development of activity theory, Engeström (2000) suggests that introduction of an artifact can serve as an interruption to routines and thus disturb the equilibrium of organizations. According to Kajamaa (2012), the role of an artifact (a concrete item or less directly observable cognitive tool such as an analytical model) in mediating behaviour is central the way people work together to (attempt to) solve problems. Because collective activity (i.e. working together in an organization) has been described by Blackler (2011) as both an enabler and producer of power, Bechky (2003) suggested that artifacts are a potentially mechanism to examine and thus disrupt the way power plays out in organizations. However, Nicolini et al. (2012) suggest that not just any artifact will do as a disruptor. To be effective in solving problems, Bechky (2003) argues that artifacts must be concrete and at the same time, loosely enough defined to be usable by all involved in the activity. Nicolini et al. (2012) recommend artifacts that pose a challenge to increase the chances that people will engage with them. They further advise that it is not always possible to predict what might be an effective artifact at any given point in time. They report that an object that is initially disruptive can become accepted and by the same token, the role of an everyday item can become an artifact that creates disruption and thus fuels change. The literature about artifacts therefore suggests that the mundane materials and processes of daily work might serve the same purpose as a more dramatic crisis in terms of disrupting routines in support of change.

Of interest to this research is the potential role of measurement as a disrupting force. The Associates for Process Improvement (API) identified measurement as the key enabler for quality improvement (API, 2016) in their “Model for Improvement”. This model (which is

foundational to quality improvement science and the Institute of Healthcare Improvement) (IHI, 2016) suggests that measurement helps identify where improvement is needed (i.e. gaps in performance) and monitors the impact of any changes to ensure that whatever action is taken leads to better (not worse) outcomes. Measurement tools act as artifacts when they focus attention on gaps and suboptimal performance. The mini-crises that result from this attention can disrupt existing practice. Measurement has also been described by Cennamo et al. (2009) as a mechanism to reduce causal ambiguity, making it easier to see the difference between what is, and what you think, is happening. They argue that measurement can therefore expose assumptions about how well the organization is doing. As Van de Ven & Poole (1995) observed, exposing assumptions, or even just attempting to do so, can generate internal and/or interpersonal conflict. Schein (1999) went further to suggest that evidence contradicting our assumptions (such as that generated by measurement) creates fear of loss of esteem, which can be disruptive either by triggering defensive routines or by encouraging second-order learning, if the anxiety is managed well. Measurement can therefore be disruptive to the equilibrium of an organization first by illustrating the need for change existing practice and by creating fear of embarrassment. Based on the above, disruption plays a key role in how continuous change occurs. While it is beyond the scope of this review to summarize all potential disrupting forces in organization, the examples of artifacts and measurement are highlighted for their relevance in the proposed research exploring quality improvement in primary care. The roles of artifacts and measurement as disruptors are key concepts guiding this action research study (see research framework at end of Chapter). The next section describes the roles of and connections between action and learning in continuous change.

Action and Learning

As described above, conversations have a role in driving actions. Actions also have a role in driving conversation. Weick (1988, p. 307) observed that “action precedes cognition”, a reversal of the Knowledge-Attitudes-Beliefs theory that was (and possibly still is) the prevailing theory for behaviour change for many years. According to Weick (1988), the role of action is to learn, not the other way around. Learning has been described as part of acting i.e. it is the process of acting, not just the outcome of the action that is important (Aronson, 1968; Salancik, 1977). Weick (1988) observed that it is often impossible to know what the right thing to do is until you try something and see what happens. However, many authors (among them, Argyris, (1999) and Dooley, (1997)) note that this learning process is difficult. Smith & Berg (1987, in Ford & Ford, 1994) observed that unless they are forced to, people may not be aware of their assumptions and therefore are much less able to question them or consider alternatives. Habits, particularly frequent, apparently successful and rewarded habits, have been observed to drive out deep problem-solving (Edmondson et al., 2001). The net effect is that intentional experimentation becomes less common over time and there is a tendency for first-order learning to predominate over second-order learning, especially in the absence of threat or crisis (Dooley 1997). However, Weick (1988) suggested that acting can support learning by conveying a sense of control, at the very least because it gives participants the opportunity to make their own sense of the results of the action. To avoid making the situation worse by acting, Weick (1988) advocates for small actions to achieve the “delicate trade-off between dangerous action which produces understanding and safe inaction which produces confusion” (p.305). Schein (1999) agreed that actions need to be small to minimize the risk and associated “learning anxiety” that impedes further action. The smallness of actions is crucial because, as Mellahi & Wilkinson

(2010) describe, it is human nature to be more forgiving of sins of omission than commission. Therefore, actions can generate conversation that can help support change, provided the actions are small enough not to scare us back into what Mellahi & Wilkinson, (2010) observed is our tendency to not-act. The role of action in supporting change is another of the key concepts guiding this action research study (see research framework at end of Chapter). The final section in this description of continuous change addresses the role of organizational power.

Power

Looking at change as continuous leads to a different understanding of how power-brokers react to and influence change than they might when considering change as a planned event. A continuous view considers that change in organizations may be affected by how those in charge respond to the disruption caused by crisis. Schein (1999) observed that crises can drive change and inhibit it at the same time. He reported that they do this by creating the necessary “survival anxiety” (p.60) to convince people of the need to change. Unfortunately, he observed that survival anxiety is often matched by equal and opposite “learning anxiety” (p.60), which reinforces defensive routines to protect from the potential loss of self-esteem inherent in the learning and changing process. While increasing survival anxiety is useful, Hanna (2008) advises that decreasing learning anxiety is just as (or even more) important as increasing survival anxiety (Smith & Elliott, 2007). Aronson (1992) noted that decreasing learning anxiety helps avoid the level of cognitive dissonance and psychological pain that Armenakis & Bedeian (1999) noted could stymy change. Schein (1999) suggests that those looking to promote change through disruption need to be prepared to invest in the development of psychological safety for those affected by the disruption. In the absence of this safety, Edmondson (2003) and Platt (1973) observed that people will invariably do what

they perceive is good for them, even if they accept that this might not be good for the organization. If those in power are experiencing too much learning anxiety, Armenakis & Bedeian (1999) observed that they may well inhibit change in the organization. Thus, as in a view of change as a planned event, consideration of change as continuous can be affected by the pre-existing distribution of power in the organization.

Another way that the lens of change as a continuous phenomenon affects the understanding of the impact of power-brokers is through their role in delaying action. As noted above, action is a very important tool in continuous change. However, despite its usefulness, Turner (1976) noted that action can be delayed by the investment of decision-makers in the status quo. Turner argues that organizations that are distanced from their stakeholders tend to develop an immunity to negative feedback and instead build a self-reinforcing culture of believing that everything is going well and there is no need to change. Further, Turner argues that attention to a well-defined, high-profile, but managed, problem can also delay action by distracting attention from a less well-defined, lower-profile problem. He notes that when powerful members of the organization (vs other members of the organization) dismiss a potential threat, particularly one that is poorly defined, there is an even greater chance of delay in action. Turner suggests that paying attention to the reception offered to the doubting Thomas who raises concerns about threats that powerful members are complacent about can be a clue about the organization's level of risk for delayed action and thus missed opportunities for change. Given the importance of action in understanding change as a continuous event, it appears that the response of those with power in accommodating dissenting voices (or not) is one way they can affect how change plays out in organizations.

The lens of change as a continuous event suggest that a third potential impact of the distribution of power in organizations on the experience of change is the relative preference of power-brokers for a slow-and-steady approach. Ford & Ford (1994) point to the butterfly effect as evidence that unplanned changes in organizations can occur quickly. Dooley (1997) concluded that ongoing change in complex adaptive systems could just as likely be quick as slow. Certainly, crises, one of the more extreme examples of emergent change, can appear to occur very quickly. Carmeli & Schaubroeck (2008) describe situations in which managers fear that change in their divisions implies that they have not been effective in preventing problems. Rapid change might not be attractive to such managers. The importance of crises (with their attendant requirement for prompt response) in a view of change as continuous suggests that the degree of power held by managers not inclined towards rapid change can affect how change unfolds in the organization.

Finally, Tsoukas & Chia (2002) observe that a world-view of change as continuous allows one to see the unfolding process and emerging drivers of organizational behavior. Cennamo et al. (2009) reports that such clarity is not always preferred over ambiguity. In fact, they report that ambiguity surrounding the cause and/or outcomes of changes can serve to advance individual interests of managers in the organization. When those holding power in the organization are better served by lack of clarity between action and outcome, there may be less support for change that might shed light on those relationships. The view of change as continuous therefore considers the extent to which power-brokers in the organization prize clarity a factor in how change plays out in the organization.

Summary of a continuous view of change

In summary, a view of change as a continuous phenomenon recognizes that change is continually emerging from the ongoing coping of *all* members of the organization, not just those in power. In this way, a view of change as continuous holds promise for transformation through the actions of all members of the organization. This view of change suggests that paying attention to conversations that emerge from the small daily actions will make it possible to increase the potential for the exposure of assumptions and concomitant learning that are necessary for positive change. Moreover, it may be possible to encourage and direct those conversations by creating crises through introduction of artifacts or new measures of existing activities. These in turn can reframe or even challenge norms and existing routines. The theory of continuous change recognizes the influence of power as *one* of the factors contributing to the equilibrium of the organization, not the *only* driver of behaviour. In so doing, the understanding of change as continuous contains the possibility of showing how change can be achieved not just by those in power, but also by those who may traditionally have been seen by themselves, or others, as less powerful in the organization. It also shows how small events can accelerate change beyond their apparently limited scope by increasing awareness of tensions between current assumptions and desired outcomes.

Conclusions

The literature shows that there are different implications for efforts to support organizational change, depending on one's understanding of how change happens – i.e. episodic or continuous. My view of change in my thesis research is based on my understanding of my role as a scholar-practitioner in my organization and the problems my organization is trying to address.

A view of change as continuous is consistent with my operational mandate as a scholar-practitioner. I am charged by my organization to equip members to measure, demonstrate and improve the quality of care they provide. This is my responsibility even though there is not yet a clearly defined framework or even consensus for measuring what quality is in primary care. Approaching this from the perspective of change as continuous builds my capacity for the deep learning necessary to make my actions as safe as possible in such an uncertain environment. Framing action as an avenue for learning rather than the final stage in implementing someone else's vision increases my ability to instil confidence in the members, increasing their actual and perceived control members have over how their work is measured and represented in their world. The intentional focus of continuous change on learning is therefore consistent with my operational responsibility as well as with my personal goals as a scholar.

Considering change as a continuous phenomenon represents an alternative and complementary approach to the problems my organization is trying to address. These problems have arisen and persist in a world where the dominant view of change is episodic. The persistence of these problems may well be influenced by factors beyond the fundamental understanding of change. The fact that others are simultaneously trying to address these problems even as I approach them by challenging the foundational assumptions of change underlines the importance of understanding change as always/already happening. This is emblematic of a common complaint from primary care providers that there is too much change. This concept is embedded in the narrative to the point of acquiring its own terminology: "change fatigue" (Perlman, 2011). Nonetheless, change continues apace. For example, Ontario has just launched yet another primary care transformation initiative called "Patients First" (Ontario, 2016). A world view of change as continuous may help navigate

the waves of episodic change that show no signs of abatement. It makes it possible to position D2D as an artifact, rather than a solution. This makes D2D a more gentle disruption than a fully-formed initiative requiring buy-in to proceed. As an artifact, D2D gives my organization a focus for conversations, making it easier to get started with small actions than might be the case in a planned change model which is commonly oriented around implementing a solution that has already been defined and therefore might feel imposed. The theory of continuous change can therefore help build capacity for what is perceived as relentless change, even among those who see these changes as planned episodic events.

Considering change as a continuous phenomenon may also help us achieve success, given the distribution of power in the healthcare sector. While primary care has been demonstrated and celebrated as the foundation of a sustainable healthcare system (Starfield, 2009), it remains woefully under-resourced compared to other aspects of the healthcare system. A worldview of continuous change can help my organization build its power in the absence of the monetary resources to which power is usually ascribed. Approaching the problem of low participation in measurement from the perspective of continuous change allows us to embrace the wisdom of the field, transforming the questions and concerns of those involved from resistance that must be overcome to the power of front-line engagement. A world view of continuous change is therefore consistent with the goals of my organization and the context in which we are situated. Rather than distracting us into an argument about which is the right or correct view of change, this approach will help my organization move forward with the work they have set out to do: improve the quality of primary care for the residents of Ontario. This pragmatic focus on moving forward rather than engaging in unproductive argument also informed my choice regarding issues of power. I chose not

overtly or directly engage in questions of power distribution as either scholar or practitioner. This is consistent with a view of change as continuous which holds that power is *one* enabler of change but not the only (or even most important) one. I was (and am) involved in the changes in my organization and therefore affect and am affected by the distribution of power. Nevertheless, I did not want to give power dynamics more power than they inherently have. Instead I set out to do for myself as I was hoping to do for others: find ways to get started. I fully expected that power (among other forces) would affect my actions and those of the members of my organization. The cyclical design of this research which incorporated deliberate phases of reflection was intended to expose factors affecting actions, including but not limited to power.

My review of the literature highlighted key concepts about continuous change that helped me reframe my problem and thus make it easier to examine. These concepts did not answer my research question. Rather, they helped me clarify a focus and approach. In the same way that transforming anxiety into fear of a particular threat can better support change (Pauchant & Mitroff, 1988), disaggregating my problem into a series of discrete events or activities made my review of what otherwise might have been an overwhelming abundance of data more productive. The scaffolding of concepts that emerged to support my examination of my research question is summarized in the following section.

Implications of Literature Review on Research Design

I assembled key concepts from my review of the literature into a framework of inter-related elements to guide my learning in this action research project. The specific elements and their relationships with each other (represented by the numbered arrows on Figure 2-3)

are described below. The fundamental premise of the theory of continuous change is that it is “always, already happening” (Tsoukas & Chia, 2002) and is the sum of multiple coping strategies by multiple players, each making their own sense of the world around them (Orlikowski, 1996). In the same way, the concepts guiding this research come from many different theories by many different authors, in my attempt to make sense of the complex experience of changing performance measurement by getting started.

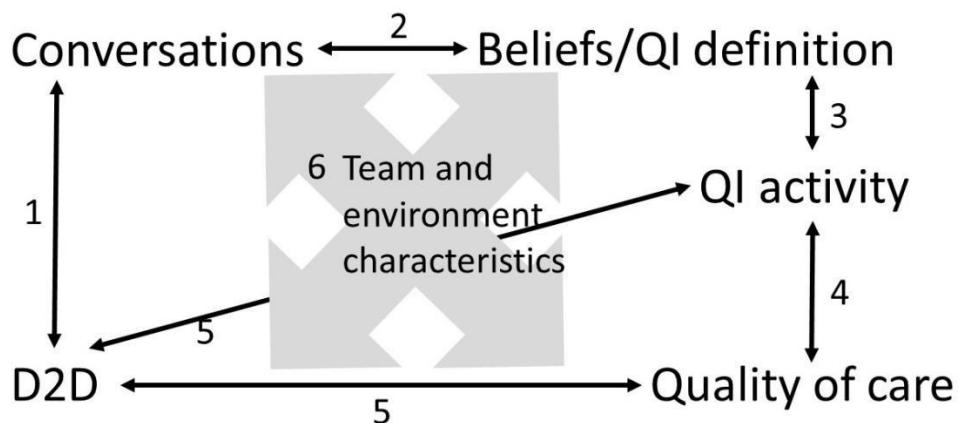


Figure 2-3: Literature-based research framework to guide data extraction for this action research project

Arrow 1: One key concept about continuous change is that, while change is “always, already happening” (Tsoukas & Chia, 2002), it can be precipitated by a crisis. An artifact or boundary object (Bechky, 2003) represents a gentle form of crisis that can sufficiently disrupt routines and create “far from equilibrium” conditions (Dooley, 1997, p. 79) in which positive change is more possible. The artifact I introduced in this study is D2D. As described earlier, D2D was a novel approach to measurement. Measurement tends to highlight the gaps between what we believe and what is actually happening (Cennamo et

al., 2009). Measurement can expose unexamined assumptions (Van de Ven & Poole, 1995), which in turn can generate “learning anxiety” (Schein, 1999) and thus affects the balance of power in the environment. Arrow 1 is also informed by the socially constructed theory of knowledge described by Weick (1988), which suggests that disruption can be expected to generate conversations. Finally, arrow 1 reflects the concept that conversations can affect power distribution by contributing to learning and change (Weick, 1988). Hence the bidirectional arrow between D2D and conversations.

Arrow 2: Another key concept from the literature was the observation of Ford & Ford (1995) that organizational changes are both reflected in, and driven by, changes in conversations. In addition, Macpherson et al. (2006) and Capelli & Smithies (2009) describe conversation as an intervention in its own right. The double-headed arrow between conversations and beliefs reflects another key concept from the literature related to Weick’s (1988) observations that while conversations can help people make sense of their environments, changes in understanding can generate further conversation. Therefore, this action research examined the extent to which conversations occurred as well as the extent to which they generated changes in beliefs and further dialogue.

Arrow 3: The D2D initiative recognized that conversations and beliefs were not sufficient to actually change performance or even behaviour. Another key concept emerging from literature about continuous change is the idea that not all conversations lead to action (Dervitsiotis, 2002). This is consistent with the observations of Taylor et al., (2006), who noted that behaviour change models that assume rational decision-making based on knowledge have not been shown to be highly effective in encouraging behaviour change or improvements in outcomes. For these reasons, I was attentive to what behaviours or activities, if any, might

follow from changes in conversation and beliefs. The double-headed arrow was based on the concept of effectuation described by Aronson, 1968; Salancik, 1977, Weick, 1988 (among many others) which holds that action affects beliefs.

Arrow 4: The concept reflected in Arrow 4 is the emerging evidence about the inconsistencies in impact of participation in QI activities on outcomes (described Hall (2014) and (Qaseem et al., (2014) in the review of QI literature in the introductory chapter). For that reason, I was also very attentive to the idea of the outcomes of QI activities (that is, actual performance on outcome measures), not just the activities themselves. The concept driving the reverse of Arrow 4 is based on the work of Edmondson et al. (2001) which described the negative impact of success on engagement in change. The double-headed arrow signifies my interest in how activities affect outcomes as well as how outcomes affect participation in activities.

Arrow 5: Given the constructionist lens of this study, and observations about the disconnect between performance measurement and outcomes (De Vries et al., 2014), I did not expect that D2D would affect either QI activity or performance directly (that is, without being mediated by changes in conversation and beliefs/attitudes). However, I deliberately looked for indications of a direct impact of D2D on performance. This would serve as disconfirmatory evidence to our fundamental assumption about the socially constructed nature of impact of D2D. In this way, I intended to reduce the refutability and thus increase the robustness of this research (Creswell, 2007).

Arrow 6: My attention to the impact of characteristics of the members of the association on any of the activities described above was driven by my knowledge of the association's needs. Operationally, it was important to consider differences between teams in size, physical design, team climate, physician engagement, rurality, EMR maturity (as measured by several

surrogates), co-location of team members and QIDS specialist connection, among other contextual features that emerged from the analysis.

The application of this framework to the cycles of data capture, analysis and subsequent interventions are outlined in the next chapter on Methodology.

Chapter 3 Methodology

General approach

The aim of this research study is to increase engagement and participation of front-line providers in meaningful measurement of performance of primary care through the introduction of D2D, a novel way of measuring quality and demonstrating value. Our problem is perfectly suited to an action research approach. Since we are bound to continue “doing” in that it is AFHTO’s strategic priority to measure, action research is an approach that helps us learn as we go, rather than taking time out for a research study. This is what O’Brien (2001) notes as the essence of action research:

“a group of people identify a problem, do something to resolve it, see how successful their efforts were, and if not satisfied, try again”

Action research is equally committed to improving outcomes in an operational activity as it is to generating knowledge from that (O’Brien, 2001; Patton, 2011). Very simply, action research is a way of “*allowing everyday people to discover their talents and shape their own future*” (Dick 2009, p. 425). In this way, the knowledge generated through action research is phronetic: it is derived from practice and thus is based on the wisdom of the field. This requires turning the “researched” into researchers through deep participation in the research process. The deep involvement of participants in the research has an operational benefit in that people are more likely to act on strategies they themselves have devised than to adopt externally imposed recommendations (O’Brien, 2001; Patton, 2011). In addition to the focus on phronetic knowledge, action research is also characterized by intentionally rejecting the myth of researcher neutrality (Johnson & Duberley, 2003; O’Brien, 2001). Since I, as the action researcher in this project, am an integral part of the action being studied, I have no

intent or illusions about being neutral. It is my operational responsibility to lead the measurement work of AFHTO and to advocate for it. I am highly invested in its success and therefore highly engaged in learning from our collective experience to constantly make it better.

Action research does not prescribe a method. As described by Dick (2009), Heller (2004) and Stern (in Patton, 2011), it welcomes any methodologies that focus simultaneously on action (or change) and on increasing understanding about the action (that is, research). According to these authors, action research has a cyclical nature, moving between phases that involve action and then phases that involve critical reflection on that action as the research unfolds. In contrast to traditional research, action research involves continuous evolution of methods in response to learning from preceding action and reflection cycles. Since the goal of this action research is to learn what is working with measurement in primary care in order to make it work better, the methodology I chose was that of developmental evaluation.

Patton (2011) observed that action research and developmental evaluation are compatible and mutually reinforcing. Developmental evaluation examines how well the program is doing what it intended to do as well as whether the program is worth doing – that is, is it still the right approach (Greenwood & Levin, 2007). Developmental evaluation is an ideal methodology when there is not yet any proven best practice to test in a subject area. It is the most suitable approach to evaluation in early stage social innovations such as this because it supports the process of innovation in ways that enable learning and adaptability (Gamble, 2008). Greenwood & Levin (2007) emphasize active participation of those involved in the program as a key driver of the utility of developmental evaluations. Stern (in Patton, 2011) positions action research as a tool for developmental evaluation in that the judgment inherent

in evaluation depends on robust analysis (that is, research) and development to further improve effectiveness depends on action. In Stern's view, an action research approach to evaluation positions the analysis aspect of research as a support for action, not just for understanding or knowledge generation. The action research approach to evaluation looks to produce knowledge through action and does so from a local well-described (vs generic) context.

Finally, Stern's view of evaluation informed by action research includes intentional blurring of the boundaries between the evaluator and the researched, resulting in deeper relationships and more equality between researchers and those who are traditionally seen as research subjects.

One risk inherent in using a participatory approach to both develop and evaluate AFHTO's measurement initiative is the loss of critical perspective. To manage this, I oriented the findings of my action research project around the critical reflection framework defined by Rolfe et al. (2001): "What, So What, Now What". These questions served as an organizing principle as well as an ongoing prod to examine the emerging data in a critical light.

Another potential criticism of my action research approach to evaluation in this project is the perception of bias. Traditional audit-style evaluations prize separation between an external "objective" evaluator and the subjects of the evaluation. This is commonly seen as a means to eliminate bias and thus more accurately describe the value of an initiative. However, the opposite may be true. Blumer (in Cook, 2006) warns that the distance creates the risk of "the worse kind of subjectivism" in which the so-called unbiased external observer fills in the gap left by ignoring the experience of participants with his/her own interpretation without recognizing that this results in the introduction of the external observer's own bias. Rather than accepting that evaluations based on participant data are biased, Cook (2006) proposes that if the participants in the initiative are not involved in the evaluation, it cannot be

considered reliable. Somekh (in Cook, 2006) labels knowledge generated apart from the participants as only partial knowledge that can over-simplify complex phenomena and thus lead to unhelpful conclusions. From a purely pragmatic perspective, traditional “objective” evaluations have often suffered from poor user acceptance and uptake of recommendations (van Winkelen, 2016). Developmental evaluation is no more successful than any other method in eliminating bias. The difference is that developmental evaluation approaches don’t even try. Instead, developmental evaluations embrace context (in which biases are well-described) and are intentionally participatory to support usability and organizational learning (van Winkelen, 2016), which are shared goals with traditional “objective” evaluations.

The methods used in this research therefore are guided by the principles of action research and conform to the general form of developmental evaluation, as described by Gamble (2008). The remainder of this chapter describes the setting and the phases of this project which unfold in an “action research spiral” (Kemmis & McTaggart, 1988). This chapter concludes with an overview the data capture and analytic approaches in this project.

Setting

The setting for this research was the Association of Family Health Teams of Ontario (AFHTO), a voluntary collective of 184 primary care teams providing care for 25% of Ontarians. AFHTO has a staff of approximately 10 people who support the membership with secretariat services and two strategic programs, one of which is the Quality Improvement and Decision Support (QIDS) program. The history and structure of the association and the program have been described in more detail in Chapter 1. The study explores what happens when a novel approach to measuring quality/demonstrating value is introduced in my

organization and examines what it takes to continually increase participation in measuring and improving primary care.

Phases of the study

The first step in a developmental evaluation is the description of stakeholders. The next steps (which are followed in each of the action-learning cycles of the action research project) are data collection (analogous to the “action” aspect of action research), framing and reporting the findings (analogous to the “learning” cycle in action research) and subsequently developing strategy and refining indicators in preparation for data collection in the next “action” cycle. For easier, more concise reference, these three steps can also be encapsulated into the 3 questions outlined in the critical reflection framework proposed by Rolfe et al. (2001). The questions in this framework, based on Argyris’ “Ladder of inference”, are “What” (that is, data collection), “So What” (that is, reflection, framing, reporting) and “Now what” (developing strategy, refining indicators). The action research project is composed of 5 phases (see Table 3-1).

Table 3-1: Phases of the action research study

Chapter	Phase	action research element	timeline	Data source
4	1: Launch of D2D	Introduction of Artifact	D2D 1.0-3.0	Operational data (web posts, announcements, minutes)

5	2: Implementation of first 3 iterations (“What”)	Action	D2D 1.0-3.0	Operational data (e.g. surveys, D2D results) and conversations (emails)
6	3: Reflections on and learning from first 3 iterations (“So what” and “Now what”)	Learning	after D2D 3.0, before D2D 4.0	n/a
7	4: Implementation of 4 th iteration (“What”)	Action	Peri- D2D 4.0	Operational data (e.g. surveys, D2D results, minutes and action logs) and conversations (emails)
8	5: Reflections on 4 th iteration (“So what”)	Learning	After D2D 4.0 release	n/a

The first phase describes D2D as the artifact and focal element of the action research. It is followed by the second phase which describes the experience of AFHTO members with the artifact, in this case, the first 3 iterations of D2D. Phase 2 is the first “action” cycle of this action research project. The third phase reflects on the experience in the context of published knowledge. It is the first “learning” phase of the project, intended to compile learnings and guide the implementation of the next (4th) iteration of D2D. The fourth phase describes the experience with the D2D 4.0 and represents the second action phase in the action research project. The final phase of this action research project is the second and last learning phase of

the project. It constitutes a reflection on the experience of implementing the 4th iteration of D2D. Altogether, this thesis describes approximately one and a half action research cycles starting at the observation point in the first cycle, moving through reflection and onto planning, action, observation and reflection on the second cycle (see Figure 3-1). The boundaries between action research cycles are less distinct in actuality than represented in these phases. The “learning” phases of the project involved action and the “action” phases of the project also involved “learning”. The boundaries are also artificial with respect to the endpoint of D2D because the initiative continues to operate and evolve even though the action research project has concluded. Nonetheless, the phases are described here with relatively clear boundaries for ease in tracking progress through and reporting on the action research project.

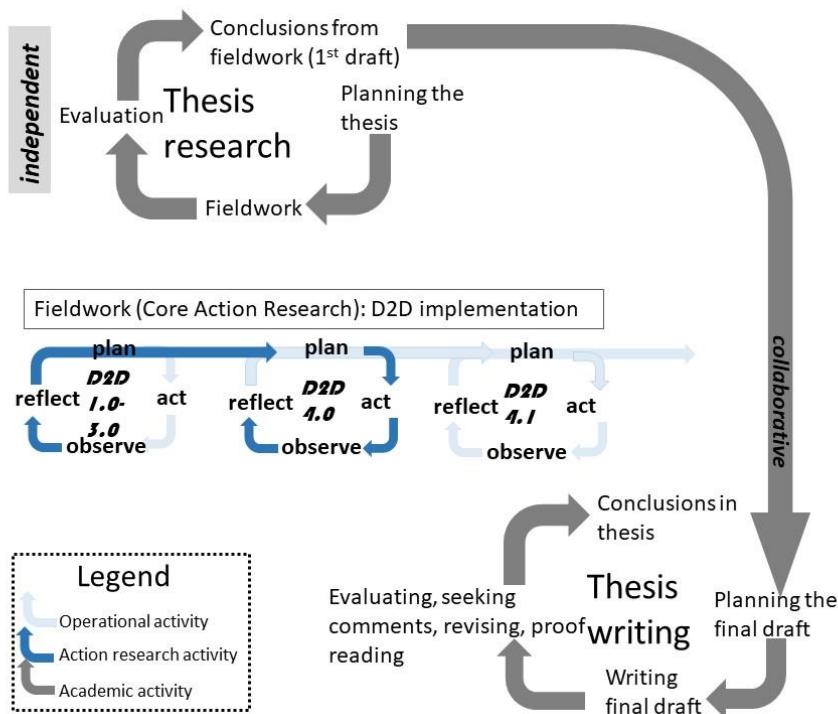


Figure 3-1: Situation of action research cycles in the context of operational activity of my organization and academic activity of my thesis (Adapted from Perry & Zuber-Skerritt, 1994)

Methods

This study is based on qualitative data and thus depends on qualitative data capture and analysis techniques. Data capture was oriented around the research framework emerging from the literature (Figure 2-3). Specifically, data were sought to describe the artifact (D2D), the nature and frequency of conversations, beliefs and attitudes about quality improvement, participation in QI activities including measurement and performance on measures of quality. Data capture processes were also attentive to the relationships between each of these topics or activities, as suggested by the double-headed arrows in Figure 2-3: Literature-based research framework to guide data extraction for this action research project. The sources of data and focus of analysis changes over the phases of the research as is expected in an action research study (King, 2004; O'Brien, 2001). As with the findings emerging from this study, the methods are presented according to the elements of the critical reflection framework defined by Rolfe et al. (2001): What? So What? and Now What?

“What?”: Data collection

Stakeholder identification

Because this was a phenomenological study, input from a wide range of stakeholders involved in the phenomenon was needed. The stakeholders involved in this study included:

- Patients: AFHTO staff have an ongoing relationship with Patients Canada, a patients' advocacy organization. In addition, AFHTO members are developing more local relationships with their patients as part of efforts to improve performance.
- Executive Directors: Each of the nearly 200-member organizations in AFHTO has an Executive Director. EDs are the single primary contact point with member organizations

and are the main administrative decision-maker in their organizations. They are paired with a Medical Lead, who facilitates relationships with physicians on the teams.

- QIDS Specialists: There are 35 QIDS Specialists deployed among 150 teams across the province. Since October 2013, they have been meeting weekly with AFHTO staff and each other and thus have developed a strong community of practice. They are widely seen as crucial enablers and change agents in measurement and QI in the teams.
- Physicians: There are about 2000 physicians working with teams. They are very difficult for AFHTO staff to reach as there is no distribution list of these people, sometimes even with their own teams.
- Interprofessional Health Professionals (IHP): There are approximately 2500 IHPs on staff at member organizations. They are the unique clinical difference from most other models of primary care in the province (which are primarily physician-based). There is a slowly growing distribution list for IHPs as well as 7 emerging profession-specific communities of practice.
- External partners: In its role as advocate for its members, AFHTO regularly and intentionally partners with funders, peer organizations and other policy or service agencies, many of whom make decisions that affect AFHTO members.

Data for this study were extracted from a variety of archival documents reflecting the input and perspective these stakeholders as described below. The intent of including input from a range of stakeholders in the data was to explore tensions and divergent perspectives as well as recognize where consensus was emerging, even when various stakeholders were not aware of it.

Data sources

The start date for action research was Feb 2014, just prior to the announcement of the first iteration of D2D. The end date was November 2016, shortly after the 4th iteration of D2D. All data examined in this thesis were generated within this period. See Table 3-2 for relevant timelines.

Table 3-2: Timeline of D2D events

Iteration	Announcement	Data submission open	Data submission close	launch
D2D 1.0	31-Mar-14	4-Jun-14	25-Jul-14	1-Oct-14
D2D 2.0	26-Feb-15	19-May-15	1-Jun-15	18-Jun-15
D2D 3.0	23-Jul-15	3-Dec-15	15-Jan-16	1-Feb-16
D2D 4.0	27-May-16	15-Aug-16	13-Sep-16	28-Sep-16

Data were gathered to describe the phenomenon of getting started with a novel approach to measuring quality in primary care. The specific data captured from these stakeholders (shown in

) is as follows:

Table 3-3: Data sources used to reflect stakeholder perspectives

Stakeholder	Relevant topics and documents
patients	Patients Canada survey development experience and results
EDs	minutes and materials for Executive Director meetings, notes from field visits and inquiries for support (emails) case studies, D2D-related survey results
QIDS specialists	minutes and materials for QIDS specialists' meetings, notes from field visits, in-person Knowledge Translation-Exchange sessions and inquiries for support (emails)
Physicians	minutes and materials for Physician Leadership Council meetings, notes from clinical consultation and inquiries for support (emails)
IHPs	materials for improvement sessions, focus group comments, survey results, notes from consultation sessions at conference
External partners	minutes and materials for meetings, notes regarding decisions and responses to proposals
All	Performance reported through D2D 1.0 through D2D 4.0
Me	Draft emails addressed to myself and/or my academic supervisor, which together form an ongoing (if intermittent) reflective journal (used as a data source primarily in the final chapter focussed on reflexivity)

Performance reports: Summaries of quantitative data reported in four iterations of Data to Decisions (D2D). These data were provided by members on traditional primary care

performance indicators as well as the indicators involved in the novel performance measurement approach being studied. All members were invited to contribute data.

Review of program documents: Archival documents such minutes and meeting materials, together comprised of 79,791 words in 21 documents. The data examined in this study includes discussions of the Quality steering committee (12 meetings) and the Indicators working group (25 meetings). QIDS program documents and relevant sections of minutes of AFHTO Board meetings where D2D was discussed (6 meetings) were also included.

Surveys: summaries of results of 8 surveys with qualitative data (e.g. nature of physician champions) and quantitative data (e.g. team characteristics, quality improvement activities).

Emails: Because I was deeply embedded in the organization and in daily communication with members throughout the design and implementation of the artifact, emails received by me from members provided a real-time picture of members' experience. Therefore, emails were examined as a source of information about how the organization was experiencing D2D as it unfolded. Emails sent by me were not reviewed because my perspective is covered in the meeting materials, web site posts, briefing notes and similar documents that I created.

Because of the massive volume of email data (29,710 messages), a sample was selected for inclusion in the study according the concept of "periodization" described by Rowlinson (in Cassell & Symon, 2004). Periodization recognizes that historical documents are not necessarily organized in any way other than chronologically. Periodization therefore involves breaking the life-cycle of the change being examined into specific time-bound periods and selecting documents (in this case, email messages) in each period. Selecting emails from throughout the life cycle, especially if they can be selected randomly as was the case in this study, minimizes the risk that the emails are selectively chosen to reflect what the researcher

wants to see. The concept of periodization was applied to this study as described in Box 1. The resulting sample included 1507 emails, comprising 2614 pages and about 670,000 words

Box 1: Application of periodization to select sample of emails to review:

- Compile list of all emails received between Nov 1, 2013 and Nov 30, 2016 (29,710), discarding part-month of Oct 2013.
- Assign each email a random number
- Select 3 emails per day (based on the random number) on each day in each peri-d2d event period (see table above for dates of D2D events). 13,070 emails received in periods 2 weeks before and after each event for each iteration. Of those, 1117 emails were selected. This represents 8.5% of emails received in these periods.
- Select 15 emails per month (again based on the random number) in the non-peri-D2D periods. 16640 emails were received from 2 weeks before the initial D2D event to 2 weeks after the last (Mar 17, 2014-Oct 12, 2016), not including the peri-D2D event intervals described above. Of these, 390 emails were selected, comprising about 2.3% of emails received in the period.

Data extraction

Data were extracted from the above sources through template analysis (King, 2004). Template analysis is particularly well suited for this action research study. It involves hierarchical sets of codes that can be merged into aggregate dimensions. This makes template analysis useful in larger data sets to facilitate a quick review at a high level to highlight areas for more detailed review. Another advantage is the capacity to assign text to more than one code (parallel coding). This can relieve the researcher of the cumbersome task of trying to choose the one and only “right” code. In this way, it is also especially suited for constructionist research such as this which assumes that there is no such thing as the “right” code, but rather that “there are always multiple interpretations to be made of any phenomenon” (King, 2004, p 256). Parallel coding is also particularly useful when the intent is to compare the perspectives of different stakeholders, an explicit goal for this study and phenomenological research in general. Also, parallel coding can contribute to rigor by helping manage the risk of poorly designed templates. As described by King (2004) the intersecting data can be examined to reveal opportunities to better define codes and/or aggregate data to best reflect the ideas emerging from the data. In the first phase of this action research project, Level 1 codes were defined by the research framework that emerged from the literature review (See Figure 2-3).

As the qualitative data were reviewed, Level 2 codes emerged for each Level 1 code to further organize the text into units of thought. A matrix was constructed to identify intersecting content generated by parallel coding. Special attention was paid to codes for which a high *proportion* (in addition to high *volume*) of text was coded to other codes. Less frequently used codes might otherwise be excluded from the review of intersecting data. This

would be inappropriate because the structure of template analysis precludes inferences about importance of any particular code based on the amount of text assigned to it. For example, much of the text at the code “light-hearted exchange” was also coded at “building relationships”, suggesting that light-hearted exchange was potentially part of the concept of building relationships. All Level 2 codes were considered in light of these emerging themes and further consolidated into aggregate dimensions. For example, the text coded at “research collaboration” was perceived to be more about “collaboration” and relationships with researchers and therefore part of the aggregate dimension of “relationships”. Based on the format for presenting templates in McDowell & Saunders (2010), Table 3-4 shows an example of an aggregate dimension with associated Level 1 and Level 2 codes.

Table 3-4: Example of coding template, showing aggregate dimensions, Level 1 and 2 codes based on the research framework and emerging from the data, respectively.

Aggregate Dimensions	First level codes	Second level codes (emerging from data)
help seeking	characteristics in general	Solving own problems
		Help-seeking
		Permission seeking
	conversations	member inquiries
		Scarce resources
		Sharing information
		Solutions from the field
		working together for more impact
		QIDSS helping QIDSS
	qi activities	Difficulty finding data

		Low EMR user knowledge
		QIDSS role in data management
	qi beliefs and definitions	Peer pressure drives data use

The code merging process was done through the lens of the problem being addressed by this action research project: how to increase participation of primary care providers in measuring and improving primary care. It is entirely likely that a different lens might guide the merging of codes towards different aggregate dimensions. Because this action research project involves several phases that are responsive to the immediately preceding phase, there are different coding templates for each phase. These are presented in the context of the analysis of data for each phase. The final coding templates are included in Appendices 2 and 3. The templates include the aggregate dimensions emerging from the Level 1 and 2 codes along with representative text for each.

The quantitative (statistical data from reports) and qualitative data for each action phase were analyzed separately, generating a series of observations, reflections and implications for each type of data. However, the experience with D2D is a whole, undivided phenomenon, not two separate sets of data. Consequently, all the data (quantitative and qualitative), need to be considered together to create meaningful actions to inform the next phases of the action research. The learnings from the two datasets in each action phase were combined by aligning observations from the quantitative data to aggregate dimensions identified in the qualitative data. The rationale for this order of operations was that there were dimensions of the experience that emerged from the qualitative data that were not observed in

the quantitative data. This was not surprising and is, in fact, the reason for using a broad range of data sources for this study. Nevertheless, the extent to which the quantitative data aligned with the themes emerging from the qualitative data validated the decision to focus on those themes in subsequent cycles of the action research. See Box 2 for an example of how quantitative data were aligned with aggregate dimensions emerging from the qualitative data.

Box 2: Aligning quantitative and qualitative data

Aggregate dimension emerging from qualitative data in first action phase:

Building relationships: there was evidence that there was attention to and success in building relationships as a way to advance measurement and improvement.

Descriptive statistics in quantitative data in first action phase:

Participation: The number of teams participating in D2D increased from 27 to 63%

Frequency of conversations: The proportion of teams having monthly (or more frequent) conversations about measurement with physician groups increased

Team climate: The average team climate functioning score increased especially in teams which reported having physician champions for improvement.

Alignment between qualitative and quantitative data

Increasing frequency of conversations may be contributing to building relationships.

Relationships may be important in increasing participation and also in better team functioning. Relationships with physicians (possibly via increased frequently of conversations) may also be important in increasing participation.

“So What?” -- Reflection and Learning

Following each “action” phase, there was a phase of deeper reflection on the experience with the action. The description of the experience that is presented in the action phase is essentially an answer to the question: “What happened?” This is the first of three questions in the critical reflection framework defined by Rolfe et al. (2001). The individual observations of activity in the action phase can then be considered collectively and in the context of evidence from the literature to answer the second question: “So What?”. This deeper reflection supports attempts to make meaning of the experience with D2D in a more comprehensive way. These reflections refine the themes emerging from the action phases to the point of suggesting specific next steps. This reflection process is roughly analogous to the framing and reporting functions in developmental evaluation, as noted above.

“Now What?” -- Preparation for the next actions

The culmination of the reflection process is the recommendation of actions. The actions proposed in this participatory action research were not framed as the ‘right thing to do’, but rather as actions the community could and would do. They were the “right” things to do by virtue of the will and interest in doing them. This is consistent with the concept of “workability” (Greenwood & Levin, 2007, pg. 68), which they see as an important aspect of quality in qualitative research. It is also consistent with a world view of continuous change, which suggests any action is the right action in that it contributes to the community’s knowledge and helps them make meaning in an ongoing way (Aguinaldo, 2004; Patton, 1999). It is also a pragmatic perspective. There is no practical way to know what the “right” thing is. Finding useful ways to increase participation in measurement and improvement in primary care is, in fact, the problem this research set out to address. It could be that the

actions taken in this action research help solve the problem for this community, in which case these actions might also be considered potentially “useful” for others. However, even then, they may not be the “right” things to do, as distinct from other actions that are less right or possibly even wrong. The constructionist viewpoint underpinning this research suggests that context and the community are the key factors in determining what is right in any situation, not objective, externally-defined ‘evidence’ (Aguinaldo, 2004; Cook, 2006; Patton, 1999). The “rightness” of the actions emerging from the attempts to make meaning from the experience and published evidence is therefore best estimated by the extent to which the community took action and the actions made a difference. Evaluating the extent to which action was taken and/or had an impact is the focus of the next cycle of the action research or developmental evaluation. Consequently, each cycle is focussed on its own set of indicators (that is, its own initial coding template), based on the actions expected or recommended from the previous cycle. Hence, the “Now What?” aspect of the action research project serves in this study as the “developing strategy and refining indicators” activity of developmental evaluation (Patton, 1999).

Ethics approval

Ethics approval was received from the Research Ethics Board of the University of Toronto, Ontario, one of the universities closely aligned with our organization as well as the University of Liverpool. This included explicit permission from the organization to access the operational data and conduct the research. In addition, the organization convened a research advisory group to guide this and similar research activities also being initiated on behalf of members. Experience with members in this organization is consistent with recent research about ethics oversight of QI that has illustrated that individual-level overt informed consent is actually

counter-intuitive to the goals of QI (Baily et al., 2006, Nerenz, 2009, Whicher et al., 2015). There is an assumed moral imperative among providers to continually improve the quality of the work they do (Baily et al., 2006, Kass & Pronovost, 2011, Platt et al., 2013, Sagarin et al., 2014, Yardley et al., 2014). This is particularly true if the work is supported by public resources, as is the case for primary care in Ontario. Given this perspective and the fact that D2D (the subject under investigation) was implemented under the explicit direction of the Board of the AFHTO, this study established and operated under community-level consent (Perneger, 2004) as expressed by AFHTO's research advisory group.

Chapter 4 Phase 1 of Action research project – the artifact

Preamble

This is the first of 5 chapters of results that together show how well getting started worked to help primary care providers measure their performance. Together, these 5 chapters tell the story of what D2D is, the conversations it started and what changed along the way in the conversations, quality improvement activities, primary care quality and, of course, the people themselves. These chapters also include lessons learned about what did and did not work well with these changes. The story presented in this research study *stops* after 4 rounds of D2D, but it does not *end*. The story of D2D and performance measurement continues, with everyone hopefully somewhat the wiser from having worked together to tell the story.

The artifact in this study was Data to Decisions (D2D). Consequently, this chapter provides a detailed technical summary of the initiative, building from the brief description in Chapter 1. The artifact called D2D is described as a performance report, but it is neither a single thing nor a single action. It is multi-faceted and composed of member engagement, supporting materials, content of the performance report and communication. D2D evolved as each iteration was implemented in response to feedback from members. The “tweaks” were made as part of the normal operations of a membership organization that is appropriately responsive to the expressed needs of its constituents. This evolution was intentional (as opposed to accidental or unconscious), but not formally structured or studied. For that reason, the experiences with D2D over the first 3 iterations are considered together as a single artifact. This chapter summarizes the key features of the first three iterations of D2D based on review of the operational notes, web posts and announcements about D2D from AFHTO. It is descriptive with very little of the action that generally is the focus of action research.

This is because the *action* started before the *research*. Because the action *research* starts in the middle of this ongoing action, it starts with a description of the artifact (Chapter 4) and environment as it was after the initial action (Chapter 5), rather than starting with the process of action.

I am inextricably embedded in the artifact called D2D. Many of the ideas and interventions were mine, by virtue of my operational responsibilities. However, I invariably respond to personal congratulations directed to me about D2D by reiterating that “D2D is a “*we*” thing, not a “*me*” thing”. It is not accurate for me to own the ideas and especially not the actions as mine alone because they were manifested only through the participation of others. For example, it was my idea to have a multi-stage consultation process to select indicators for D2D. However, I would not say that “I did the consultation” because it only happened through the support and active participation of AFHTO members. Also, many of the actions taken with D2D were the result of me making small, initiating moves that then became a series of events involving others. I have chosen not to privilege the initiation of the action as distinct from the actual events that end up taking place. This is mostly because I feel it is counter-productive to my operational efforts to build capacity, confidence and momentum for collective action among members. I am attentive to how my language contributes to the rebalancing of power associated with the introduction of D2D. I therefore intentionally use first-person-plural pronouns (or passive voice) to describe the measurement work of AFHTO, even if it is occasionally more aspirational than accurate. Nevertheless, I appreciate that it is useful from an academic perspective especially in an action research study to be clear about my role in the problem. To that end, my actions are presented with higher profile in this thesis than they are in the context of my role as a practitioner. I offer apologies to and beg the indulgence

of my fellow travellers on the D2D journey for occasionally usurping their contribution to the work as it is described here.

Member engagement

Governance

D2D was born out of AFHTO's strategic priority to demonstrate the value of interdisciplinary team-based primary care relative to other models of care such as family doctors working alone or in groups. (See Chapter 1 for more detailed description of teams). A formal proposal for a measurement report was approved by the AFHTO Board in Feb 2014. Informed by the literature regarding the importance of early action (Weick, 1988) but in small safe ways to learn (based on an intuitive awareness of the disruptive potential of measurement) (Cennamo et al., 2009; Schein, 1999), this proposal framed D2D as a way to 'get started' with measurement. D2D 1.0, the first iteration of the report, was released in Oct 2014. Producing D2D continues as a routine part of the work of AFHTO staff and members.

Indicator selection process:

With my help, AFHTO members selected the indicators included in D2D. I guided them through an 8-step selection process (see Figure 4-1).

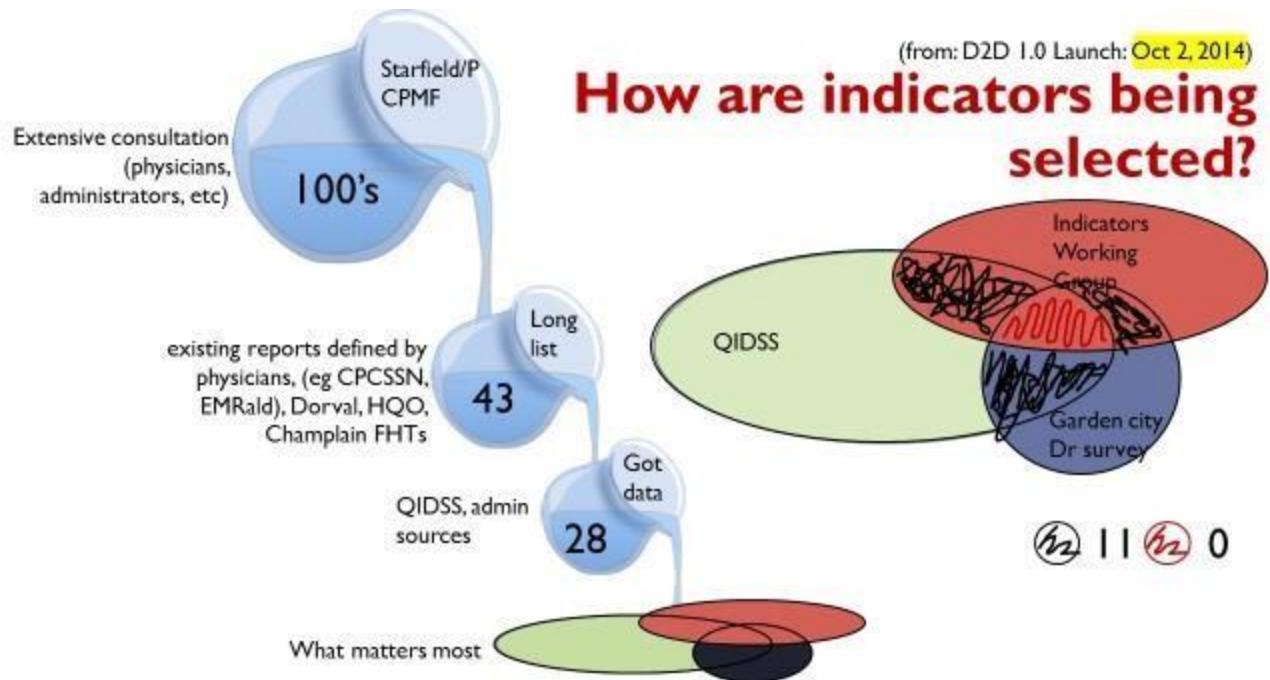


Figure 4-1: Summary of D2D indicator selection process

From the outset, I was explicit in framing D2D as being responsive to input from members regarding intentional evolution of the report. Consistent with the ideas of effectuation and sense-making described by many authors (for example, Aronson, 1968, Salancik, 1977; Weick 1988), I was intentional about creating space for the learning that I hoped would happen through the action of participating in D2D. To that end, I supported the member-based steering committee in developing and implementing a formal algorithm for consideration of new indicators. One of the strategies to ease the introduction of new indicators was to frame them as “exploratory” initially. The intent of assigning exploratory status was to encourage conversation and action at the team level to prepare for incorporation of the “new” indicator as a full-fledged performance indicator in a subsequent iteration of D2D. This process continues to guide consideration of new indicators in each iteration of D2D.

Participation

The initial definition of participation in D2D 1.0 was explicitly broad. My intent was to include as many members as possible in the initiative from the beginning and avoid labelling some members as “early adopters” and others as “laggards”. This nurturing approach is characteristic of change facilitation based on a view of change as a continuous, ongoing phenomenon (Palmer & Dunford, 2008). To that end, I initially defined the concept of “participation” to include voting on the indicators to be included in D2D, completing surveys released in concert with D2D data submission process and/or contributing data to D2D. There were several sign-up processes for D2D 1.0. My intent was to embrace the ambivalence regarding the change as is common in leadership of change as a continuous phenomenon (Piderit, 2000). I did this to allow teams the time to gradually build commitment to the project, culminating in the final voluntary act of contributing data to the report. In this way, I was trying to make the change as small and psychologically safe as possible, in keeping with the advice of Schein (1999). By D2D 3.0, the definition of participation had drifted without overt announcement to mean “contribution of data”, suggesting that members felt it was no longer (if ever?) problematic to distinguish between those contributing or not. Therefore, I used this latter definition of participation in subsequent analysis of D2D in this thesis.

Supporting materials for D2D

A guide was produced for each iteration of D2D to help teams compile their data in preparation for submission. In addition to being operationally expected, the resources were provided to make D2D easier (that is, decrease what Schein (1999) called “learning anxiety”).

Details of the supporting materials are included here to provide context for later discussions about help-seeking behaviour.

Access to external “administrative” data: The guide included instructions for accessing externally produced reports. Chief among them was administrative data such as physician billing claims and hospital admissions. Access to these data initially involved complex multi-step processes to establish permissions and mechanism for sharing anonymous data with teams. After this (and partly in response to my express request on behalf of AFHTO), these data were made available via a portal managed by the provincial health quality organization, Health Quality Ontario.

Data capture resources: Data capture resources were collaboratively developed with and disseminated to and by QIDS specialists. The resources included links to web pages with standardized computer programs to extract data from EMRs in a consistent way. These programs were developed by QIDS specialists with the interest and necessary advanced skills in technical aspects of EMRs. This enabled all providers to access the same data, independent of their own local expertise in the considerably technically demanding task of computer programming. The data extraction programs were developed for use in 3 EMR products, which covered about 85% of AFHTO members.

Data submission form with instructions: For D2D 1.0, teams entered data into a blank template distributed by email to each member of AFHTO. Participating teams then submitted completed templates to an anonymous electronic “drop box” at which point they were uploaded into the web-based report by staff of an external partner agency. Anonymity was desired by members to reduce the fear of the data being used against teams which could in turn prompt defensive reactions (Argyris, 1999) against D2D. For all subsequent iterations of

D2D, data entry was done by teams themselves directly into a web-based form. For all iterations of D2D, members were encouraged and enabled to submit only as much data as they wanted or could access. Welcoming whatever teams were able to do instead of setting standards teams claimed they could not meet is an example of reframing what might have been considered as resistance into engagement in change, another hallmark of change management from the perspective of change as a continuous phenomenon.

Content of report

The D2D report has evolved into three main areas of content. The initial iteration of D2D had a single unit of content: performance on all the indicators in the report. It continued in subsequent iterations but was joined by sections for cost, quality roll-up and submission of data for expanded indicators. These sections are explained in more detail below.

Performance indicators

The multi-stage consultation process to select indicators resulted in a short list of 11 performance indicators for D2D 1.0 (see Table 4-1). As described above, the sources of data for these indicators were administrative data (such as billing or hospitalization records), patient experience surveys conducted by members and EMR data extracted by members from their own systems (see Table 4-1). Most of these indicators were retained for subsequent iterations with some additions and subtractions as shown in Table 4-1. The process for selecting/retiring indicators is described above. After D2D 1.0, these indicators were referred to as “core” D2D indicators to distinguish them from other types of indicators being introduced in subsequent iterations.

Table 4-1: Core performance indicators in D2D iterations (see Appendix 1 for detailed definitions in Data Dictionary)

Indicator	Data source	D2D 1.0	D2D 2.0	D2D 3.0
Readmissions within 30 days of hospitalization	admin	X	X	X
Continuity of care – individual physician	admin	X	X	X
Childhood immunization	EMR	X	X	X
Availability of same/next day appointment	patients	X	X	X
Reasonable wait for appointment	patients	X	X	X
Patient involved in decisions about their care	patients	X	X	X
Colorectal cancer screening	admin	X	X	X
Cervical cancer screening	admin	X	X	X
Influenza immunization for 65+ year olds	EMR	X		
Patient has opportunity to ask questions	patients	X		
Providers spend enough time with patient	patients	X		
Courtesy of office staff	patients		X	X
Continuity of care – team of physicians	admin		X	X
Follow-up after hospitalization	EMR		X	
Diabetes management	EMR			X
Time spent delivering care	team			X

Cost:

Although included as part of the core performance indicators in D2D from the first iteration forward, cost is considered somewhat separately from the other performance indicators as more of a system outcome than a potential focus for local team-based improvement activity. It is therefore presented here as a separate section of the content of the D2D report.

I responded to guidance from research partners to include average per capita healthcare costs for each team in the first iteration of D2D and all subsequent iterations. This indicator was not novel in itself, having been developed and widely used at a regional or provincial level in policy making in government and by various researchers for several years. However, it had not previously been made available at the team level to front-line providers prior to D2D.

Quality Roll-up indicator

In response to AFHTO's endorsement of the Starfield (1998) vision of comprehensive care, I introduced the composite Quality Roll-up (QRU) measure in D2D 2.0. It was an explicit reflection of the elements that Starfield (1998) identified as important in high quality primary care: Continuity, Comprehensiveness, first-Contact access and Coordination. These principles first surfaced in AFHTO's external communication in the 2013 annual reports. They informed indicator selection, but were not highly visible in the initial D2D report. The QRU indicator was based on the pioneering work of George Southey, (Southey & Heydon, 2014) a family physician working in a primary care team in Ontario and among the founding Board members of AFHTO. I adapted it to work across multiple teams. It considers performance on a range of different primary care indicators weighted according to how important each is to patients in terms of their relationships with their provider. Southey developed and used this approach to measuring quality within his own practice for more than

12 years. Figure 4-2 shows a schematic of the process of generating scores for the QRU indicator that I adapted from the earlier Southey approach and used in D2D.

Version	Sample	Contents	Normalization	Patient priority weights	Generate composite	Share final score
Original	1 practice, 10+ years	Performance on 60 component indicators – missing values set to mean	Normalization, using thresholds based 6 physician colleagues	Weighting according to input of in-depth conversation with 6 patients	Sum of weighted, normalized performance scores	Share quality roll-up score with drill down into domains
<i>Evolutionary actions</i>		<i>Invite all teams to contribute up to 60 indicators; select 14 to yield sufficient sample size and comprehensive view of primary care, validate composite using QOF data</i>	<i>Conduct membership survey to establish consensus on thresholds then establish statistical thresholds based on data distribution</i>	<i>Conduct population based patient surveys</i>	<i>Sum of weighted, normalized performance scores</i>	<i>Display score across membership, with time trends and drill down to components</i>
Spread 1	67 teams, D2D 2.0-3.0	Performance on 14 component indicators, missing values imputed randomly	Normalization, using thresholds based on member survey	Weighting according to results of population-based patient survey V1	Sum of weighted, normalized performance scores	Final quality roll-up score, with drill down to domains
Spread 2	115 teams, D2D 4.0-5.0	Performance on 14 component indicators, missing values imputed randomly	Normalization, using thresholds based on 25 and 75 th percentiles in D2D 2.0	Weighting according to results of population-based patient survey V1	Sum of weighted, normalized performance scores	Final quality roll-up score as % of max possible score, drill down to components
Spread 3	115 teams + 5 non-team practices, D2D 6.0	Performance on +/- 14 component indicators (pending QOF analysis), missing values imputed randomly	Normalization, using thresholds based on 25 and 75 th percentiles in D2D 2.0	Weighting according to results of population-based patient survey V2	Sum of weighted, normalized performance scores	Final quality roll-up score as % of max possible score, drill down to components

Figure 4-2: Evolution of the process of generating the composite quality score

The adapted process I used to generate the QRU started with the selection of component indicators. The QRU currently includes 14 individual performance indicators. These were selected based on availability of data (that is, number of teams contributing data), minimal correlations with other indicators in the set and coverage of as many as possible of Starfield's principles (1998). The next step was the normalization of performance for each of the indicators to make comparisons between indicators with widely divergent distributions more meaningful. In the third step, the normalized performance on each indicator was weighted according to how important that indicator is to patients in their relationship with their primary care provider. The weights were derived from a purpose-specific survey of

patients asking them to consider many common primary care performance indicators and rate how much each matters to how they feel about their primary care provider. This survey was done on a population-basis separate from the local patient experience surveys conducted by each primary care team on an ongoing basis as part of their own local performance measurement activities. A fourth step was the imputation of missing data. This is necessary to be able to accommodate those teams who did not submit data for all 14 components. The final step was the summing of the weighted normalized performance of each indicator and presentation as a percentage of the total maximum score. The random nature of imputation makes it hard to interpret the meaning of the score at an individual team level. However, it confers confidence in the average score based on a large number of teams, resulting in a stable estimate of the QRU score at the membership level, even with missing data for some teams (OECD, 2008, Raghunathan, 2004). The balance between the value of the composite score at the level of association and its limited usefulness at the individual team level, both of which are the result of imputing missing values, is emblematic of the concept of equilibrium that pervades the view of change as a continuous phenomenon (Orlikowski, 1996).

Expanded indicators

The introduction of the Quality roll-up indicator made it necessary to get more data than that represented by the core D2D indicators. Therefore, starting in D2D 2.0, teams were invited to consider contributing data to an expanded set of indicators. My intent was to get as much data as possible to be considered in the generation of a composite measure of quality (see above). Over subsequent iterations, the expanded set of indicators served additional purposes of providing teams with high capacity to access data an opportunity to contribute more. This seemed to be motivating to them as expressed by their request to have summary

reports of these data. It also provided early signals of growing capacity and interest among AFHTO members to contribute data for specific topics. This was reflected in the addition of an option to the process for selecting new indicators for D2D that is, introducing new indicators via the expanded indicator set (see indicator selection process above). The reaction of members to the introduction of the composite measure contributes to the role of D2D as an artifact in this action research project.

Team context

EMR maturity:

Because primary care Electronic Medical Records (EMRs) contain data about all the healthcare events experienced by the patient, EMR data are potentially very useful for measuring and improving performance. They are also oriented around the patient, not specific bounded events like a laboratory test, a hospitalization or visit by community service provider. EMR data are considered to be near-real-time data for primary care, being updated in the course of the interactions of patients with their providers. Finally, as the record used by all providers in the team, they represent the single best source of data about *team-based* primary care, which is the focus of AFHTO's efforts to demonstrate value. Therefore, EMR data were of interest to AFHTO members. However, EMR data are widely considered to be of such poor quality they cannot be used for measurement purposes. The role of EMR data in measurement is frequently described and dismissed in 4 words: *Garbage in, Garbage out*. To move beyond this impasse, I invited AFHTO members consider trying to improve EMR data quality the same way they were trying to improve quality of care – that is, by measuring it. There are therefore two measures of EMR maturity in D2D.

Hospital-EMR integration is a categorical measure of the team's information infrastructure. It describes the existence of an automated connection between the information systems of the team and the hospital. These connections pass data such as discharge notes, consultation reports and diagnostic imaging results from the hospital systems directly and automatically into EMR systems.

The second is a quantitative measure of the quality of EMR data. The measure I initially proposed was a subjective “high-medium-low” categorical scale. This was rejected as useless for comparing between teams and driving change in data quality. In response, I proposed a composite measure based on the EMR data quality framework (Bowen & Lau, 2010). For D2D 2.0, the composite EMR data quality measure had only one component: concordance between EMR cervical cancer screening rates and rates based on the “gold standard” provincial cancer registry maintained by Cancer Care Ontario. In subsequent iterations, additional elements were added such that by D2D 3.0, the EMR data quality composite indicator included 5 components as shown in Figure 4-3. The response to the dismissal of the initial EMR data measurement scale is another example of how change management through the lens of change as a continuous phenomenon deals with what might be labelled as resistance when thinking about change as a planned event.

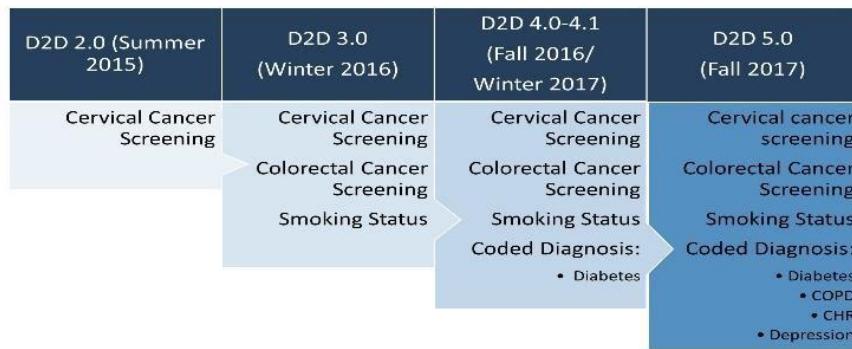


Figure 4-3: Evolution of the EMR data quality composite measure

Team setting and functioning

Three indicators were included in D2D to describe the setting of the primary care team.

These included “rurality” (rural or urban), patient panel size (<10000, 10000-30000 and > 30000) and teaching status (non-teaching, teaching and formal academic affiliation).

Additional elements to describe the team were included in a survey released at the time of submission of data to D2D. Through the survey, I sought information about factors that were thought to be enablers of QI such as access of the team to a QIDS specialists, the governance model of the team (e.g. community or provider led board), estimates of team functioning (via Team Climate Inventory), drivers for goal achievement and presence of a physician champion for QI in the team. The survey also sought evidence of QI behaviour such as conversations about improvement and performance, maturity of use of EMR and participation in data standardization efforts such as those being developed by QIDS specialists. Members were encouraged to complete the anonymous survey even if they had not been able to contribute data to D2D. They were also encouraged to provide their

anonymous team code to allow linking of the team context information in the survey to the performance data in D2D.

Communication

Data presentation web-page: I enlisted an external partner (OntarioMD) to upload and format the data presentation for D2D 1.0 at no charge. This was partly due to scarcity of skilled resources within AFHTO. It was also partly because AFHTO members were more comfortable to have this work done by an external partner to minimize the risk of AFHTO learning the identity of teams contributing data. By D2D 2.0, members agreed to have AFHTO staff manage the data submission process. From the outset, the D2D report was available to all AFHTO members, regardless of contribution of data.

Identity management: D2D 1.0 was completely anonymous. The need for anonymity was established by AFHTO members who felt it was crucial in creating the necessary psychological safety for participation. Teams could access their own data in the report by entering their own anonymous code. Because some teams lost their code and therefore access to their data, teams were invited to share their code with AFHTO staff in subsequent iterations. Teams also agreed that AFHTO staff could use the information to contact them if it turned out they were a top performer so that they could be invited to share more details about their success with peers. By 2.0, most members agreed to voluntarily identify themselves to AFHTO staff.

Internal: The release of each iteration of D2D was announced via membership-wide webinars and a membership-wide email. Communication about D2D 1.0 internally was through the usual email distribution process to Executive Directors, Physician Leads and Board Chairs of member organizations. I discussed D2D frequently on the weekly teleconference meetings with the QIDS specialists as well as at quarterly meetings of the steering committee and the

monthly meetings of the Indicator working group. These were the groups charged with overseeing the design and implementation on behalf of members. D2D was presented in one concurrent session in one stream at the annual AFHTO conference (among 45 other sessions in a total of 6 streams).

The content of communication about D2D evolved over time. Initially, there was no intent to compile a summary of D2D data for external publication. However, by version 3.0, contributing to D2D was framed as serving an intentionally externally-facing goal (that is, to help strengthen AFHTO's advocacy). The volume of communication also gradually increased. For example, D2D was mentioned 10 times in the 2014 annual report, only one of which was a section title. In contrast, D2D was mentioned 20 times in the following year's report and covered 3 of the 12 pages in the 2016 annual report. It was highlighted at subsequent annual AFHTO conferences via a booth, swag, a conference "game" in addition to another concurrent session. In addition, I introduced a bi-weekly eBulletin that focussed explicitly on news and activities related to measurement and in particular D2D. Readership of the eBulletin (for which people could enrol or unsubscribe voluntarily) quickly reached a plateau of about 40% (see Figure 4-4).

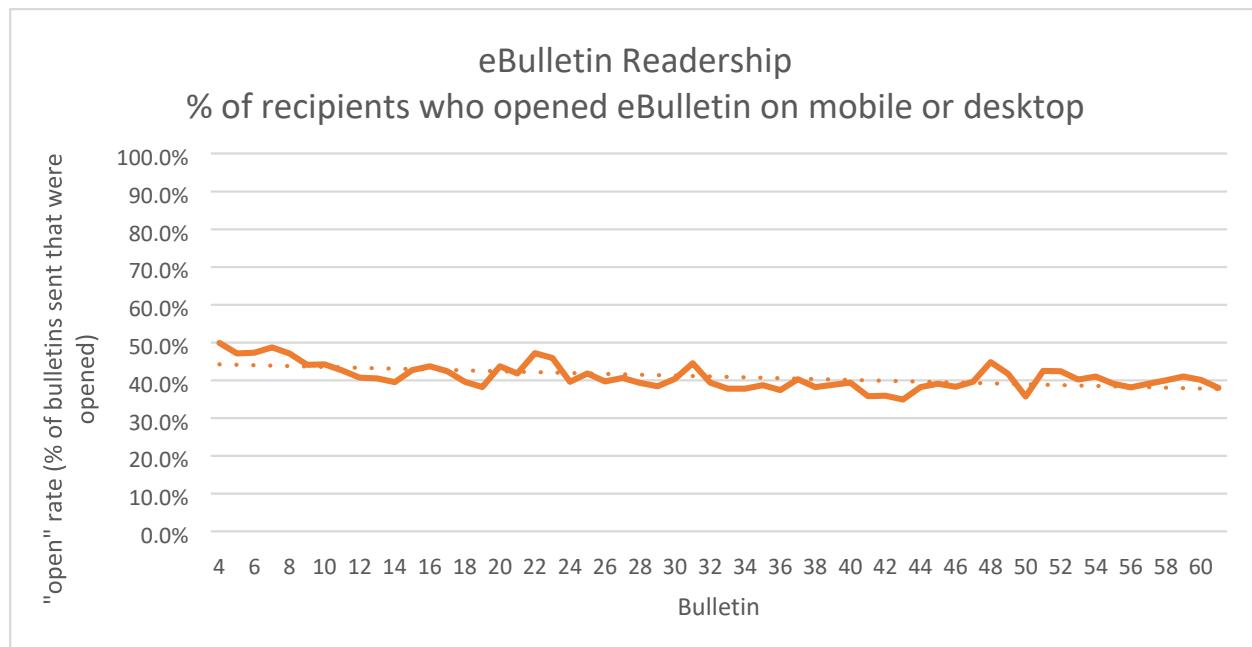


Figure 4-4: Readership of eBulletins

The D2D report was available to all members via the AFHTO members-only web page. Teams contributing data could enter their own anonymous code to refresh the data display to show their team's performance relative to peers and the entire D2D database average. Teams could choose their own peer groups based on size (number of patients), teaching status and rural or urban settings (see Figure 4-5).

Data to Decisions 1.0: Advancing Primary Care



Figure 4-5: Layout of first iteration of D2D report

While the report was intended to be used in its electronic form (to enable use of the interactive peer selection and display features and hyperlinks to point to additional information), there was considerable interest in and request for help with printing the report. The difficulty in printing the report remained an outstanding and oft-mentioned gap, or even obstacle, throughout the first 3 iterations of D2D.

External:

The external communication about the first iteration of D2D involved simply forwarding the internal announcement to 3 members of the Primary Health Care branch of the government: “*FYI – just sharing the good news that we launched the D2D 1.0 online tool today...If you’re interested [in a demo], we’ll schedule it...D2D 1.0 is a summary of primary care data that are currently available, comparable and mean the most to AFHTO members in their efforts to advance quality of care for their patients*”. The announcement described the process and rationale, but not the level of performance of teams. The announcement was not dedicated to D2D in that it included 3 other announcements relevant to members. It was followed within one month with an external release celebrating the performance of teams relative to other providers as demonstrated in an externally-produced sector-wide report, referring to “encouraging results among AFHTO members”. In subsequent iterations, there was more advanced planning of the messaging about D2D to external audiences. For example, a working group of the Board was convened to preview the results of D2D 2.0 for the purposes of preparing external messages to coincide with the release of the report of AFHTO members. The messaging was also more emphatic and dedicated exclusively to D2D (with no other topics included in the messages). The opening statements for the public announcements about D2D 2.0 and 3.0 were as follows:

“*The Association of Family Health Teams of Ontario (AFHTO) has just released **Data to Decisions (D2D) 2.0** – a ground-breaking report on performance in team-based primary care in Ontario*”.

“*Today’s release of **Data to Decisions (D2D) 3.0** demonstrates significant progress by family health teams and nurse practitioner-led clinics*”.

These messages were distributed to nearly 100 people in more than 40 organizations at the same time as the release to members. In addition to the announcements of the launch of the D2D reports, I made presentations about D2D on behalf of AFHTO at several Canadian and North American primary care conferences.

Summary of Phase 1: The artifact

D2D is commonly perceived to be a performance report. However, as described at the launch of the first report, D2D is “*the tangible result of our collective “get started” effort in team-based performance measurement*”. The D2D report itself is only one of a range of activities involved in advancing measurement. Together, this collection of activities is the artifact that serves as the focal point for this action research study. The subsequent chapters of results (each representing one phase of the research) describe the response to the artifact (Chapter 5), the incorporation of learning from these responses into subsequent iterations (Chapter 6), the experience of members after changes were implemented to D2D (Chapter 7) and the reflections on those changes (Chapter 8).

Chapter 5 Response to first 3 iterations of D2D (Phase 2 of action research study)

Preamble

This chapter summarizes the response to the first to the first three iterations (D2D 1.0 to 3.0). By the time this action research study began, these three iterations had already occurred. As described in Chapter 4 (Phase 1), the initiative evolved intentionally, but without formal study through those first 3 iterations. The response to these first 3 cycles combined is Phase 2 for this action research because it represents the beginning of formal study and reflection with the intent to contribute to practical knowledge. This is distinct from the approach to learning in earlier iterations, which was focussed more on continual improvement of an operational initiative. This action research therefore starts with description and observation of the responses to an ongoing action, rather than starting with an action as might reasonably be expected in an action research. Consequently, this chapter, like the previous one, is descriptive in nature, to set the context for the action described in Chapter 6. It also includes brief summaries of attempts to make meaning of the individual observations. The distinct areas of observation described in this chapter are: process and structure measures related to D2D, performance on D2D indicators and conversations among AFHTO staff, members and external stakeholders.

Table 5-1 is intended to make it easier to navigate through the wide range of what might appear to be the unrelated series of observations described in this chapter.

Table 5-1: Guidance for review of Chapter 5

Main section	Areas of focus of observations in each section
Process, structure and team characteristics	Summaries of quantitative data about participation in measurement, prevalence of physician champions and QIDS Specialists and team characteristics such as team climate, frequencies of conversations about measurement and access to EMR data
Performance measures	Summaries of quantitative data on measures of EMR maturity, individual D2D indicators, novel composite measure of quality, cost and finally, relationship between quality and cost
Conversations	Aggregate dimensions emerging from review of qualitative data

Process and structure

A note about data sources for processes, structures and team characteristics

The data on which this chapter is based came from several sources. The data source for participation is enrolment data for the D2D initiative. These data are relatively uncomplicated. The second data source, which was used to describe the other process and structure measures bears further discussion. These data came from surveys that AFHTO members were invited to complete at the deadline for data submission for each iteration of D2D. Analyses of survey data were incorporated into ongoing discussions with AFHTO members and oversight committees in the form of announcements, briefing notes, evaluation reports and other documents. For example, the focus on inviting members to participate was based on survey responses, but not overtly referenced as such. In addition, action was taken in response to key findings in the surveys. For example, selection of indicators was heavily

influenced by survey results. The descriptions of processes and structures related to D2D (except for participation in D2D) are based on survey data that was used by AFHTO staff in various decision-making and communication materials. There were no formal stand-alone summary reports of the results of each survey. This section is therefore based on review of the subset of survey data that was of interest to and used by AFHTO staff in subsequent communication with AFHTO members. It does not consider the entire spectrum of findings from the surveys, which was not available for this study without *de novo* analysis of individual survey responses.

Participation

Participation, defined as “contributing data to D2D”, is shown in Figure 5-1. By D2D 3.0, 63% of members contributed. Key reasons members gave for participating in D2D included: “*we were asked to*”, “*it's a valuable comparison tool*” and “*to help AFHTO help us*”. I framed the level of participation as “*nearly two thirds of members*” while staff at government chose to describe it as “*practically all AFHTO members*”, suggesting they perceived 63% to be a high rate of participation. In addition to increasing numbers of teams contributing to subsequent iterations, there was an increasing amount of data contributed by each team. For example, virtually no teams contributed data for all of the indicators for D2D 1.0. By D2D 3.0, even though I and my team provided ongoing reassurance that data submission for D2D was not an “all or nothing” exercise, over 80% of teams contributed data for all core D2D indicators.

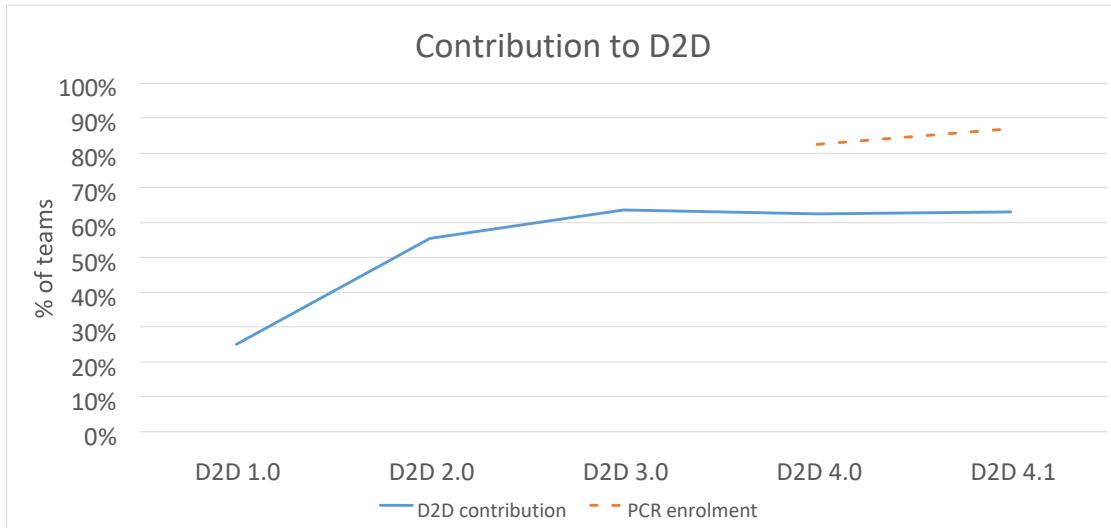


Figure 5-1: Contribution to D2D by iteration

To facilitate better access of teams to data, I advocated on behalf of AFHTO members to introduce team-level enrolment in the externally-produced Primary Care Practice Reports in addition to the existing individual-physician-level enrolment. Within weeks of being invited by AFHTO to take advantage of this new service, over 70% of teams signed up for the report, compared to less than 10% enrolment in the report by individual physicians up to that point. As I had predicted (and came close to promising when making my request!), team-level enrolment was also accompanied by the biggest increase in individual physician enrolment in the history of the report (see Figure 5-2). This suggested that inviting is an important strategy to encourage participation in measurement activity. It also suggests that it might matter that the invitation comes from a trusted party (in this case, AFHTO staff) and is directed at parties already engaged in measurement (teams rather than individual physicians, in this case). Further attempts to advance performance measurement might therefore leverage the receptiveness of teams to invitations from AFHTO staff.

Physician enrolment: Primary Care Practice Report

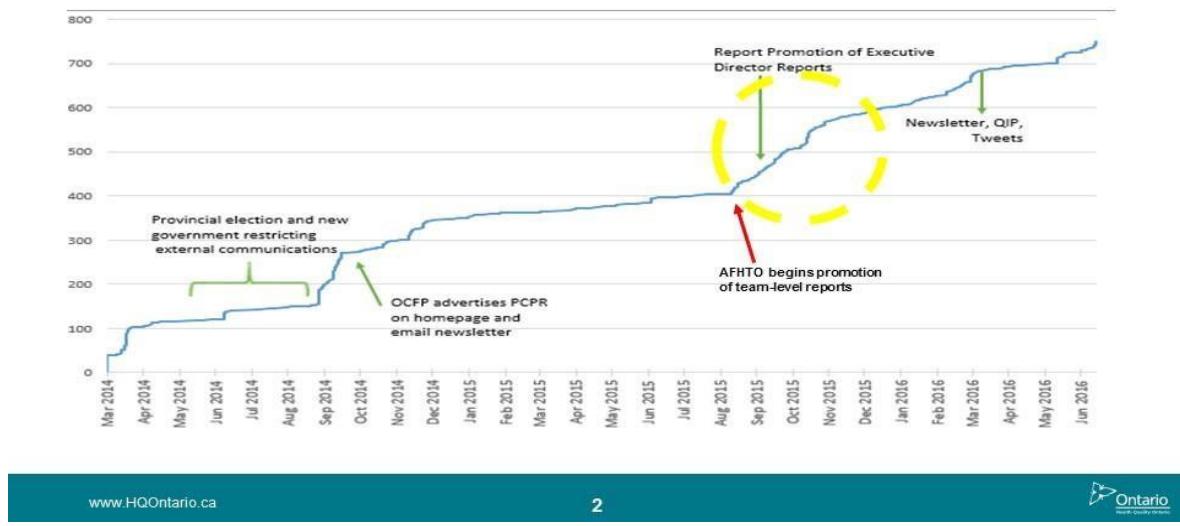


Figure 5-2: Primary Care Practice Report uptake after introduction of team-level access

Structures

Some of the structures related to D2D included enablers of contribution to D2D (e.g. physician champions, availability of QIDS Specialists), and team characteristics (e.g. team climate). These are described in more detail below.

Physician Champions

According to the surveys done in concert with D2D, 33% to 54% of teams (circa D2D 2.0 and 3.0, respectively) felt they had a physician who champions quality improvement. The most common characteristics attributed to physician champions were: a love or perceived value of data, personal interest in being better, active participation in quality improvement activities and projects, advocacy for QI to other physicians and positional authority as medical or QI lead (or similar role). Physicians expressing interest in D2D (either by participating on committees or contributing spontaneous feedback) were highly responsive and effective in building relationships with other physicians. For example, 80% of physician champions asked

to introduce AFHTO staff to a colleague for the purposes of getting broader input into D2D made the introductions. Of the physician colleagues introduced to AFHTO in this way, 100% agreed to provide input. This suggests that physician champions could be important enablers of QI.

QIDS Specialists

Most teams participating in the surveys had QIDS Specialists support, starting at 71% and reaching 88% for D2D 3.0. This is close to the overall proportion of members who have QIDS Specialists support (83%). There were many comments like “*We would not have been able to participate in D2D without the QIDS Specialists*”. The odds ratio associated with participating if a team had a QIDS Specialists was 1.6. It was not statistically significant ($p=0.32$), possibly due to response rates in surveys on which the data were based. The overall direction from the data seems to be that QIDS Specialists played a significant role in a team’s actual or perceived ability to participate in D2D.

Team characteristics

Team climate inventory

Team climate as assessed by the Team Climate Inventory (Kivimaki & Elovainio, 1999) was among the characteristics of teams explored in the D2D surveys. Team climate is a measure of the quality of interactions and effectiveness of teams. There was an increase in the team climate score (0.64 higher on scale of 1 to 5) between D2D 1.0 and 2.0. This difference was statistically significant ($p<0.05$) and suggests an improvement in team functioning over the time period. Given that the team climate inventory includes aspects of information sharing, it is reasonable to assume that the improvements in team climate are related to increased frequency of conversations and increased prevalence of physician champions.

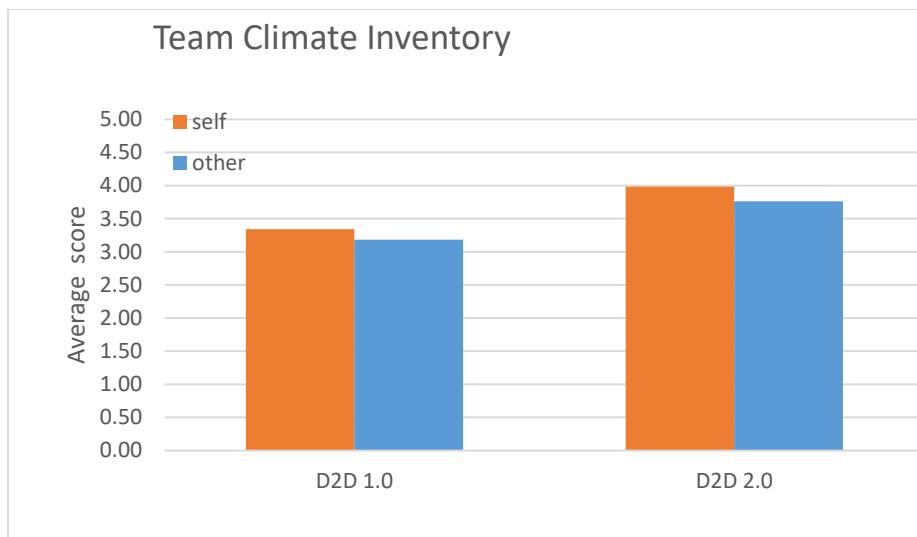


Figure 5-3: Team Climate Inventory scores

Frequency of team-level conversations about performance

Teams responding to the D2D surveys reported increasing frequency of conversations with physicians about performance with each iteration (see Figure 5-4). The differences in frequencies between iterations were not statistically significant, but were consistently higher (except for conversations about the Quality Improvement Plan (QIP), an externally-required report) over all 3 iterations. The increasing frequency of conversations suggests that D2D 1.0 was successful in achieving one of its goals, that is to prompt conversation. Since one of the characteristics of an artifact is its ability to generate dialogue (Macpherson & Jones, 2008), the increasing frequency of conversations also validates the premise of this research that D2D is serving as an artifact.

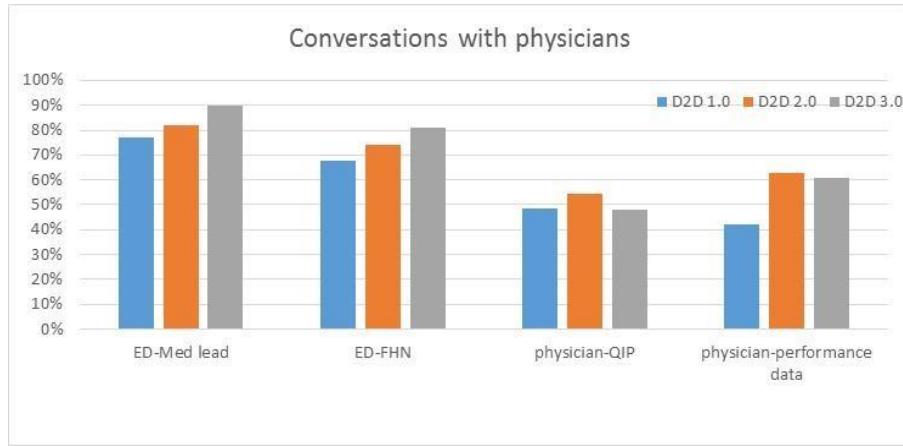


Figure 5-4: Proportion of teams with frequent (monthly or more often) conversation with physicians about performance.

Access to EMR data

Standardization for all 3 types of data involved in D2D increased in the 3rd iteration (see Figure 5-5). However, there was no increase in EMR functionality over the same time period. The increase in standardization in the absence of improved EMR functionality suggests there was increased engagement in and effectiveness of people-based processes such as policies and education related to consistent data entry, among other things.

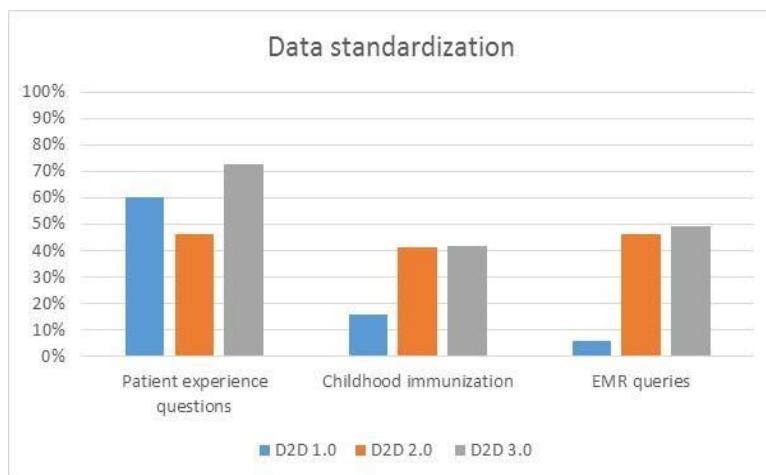


Figure 5-5: Percent of teams using standard tools and definitions to access data for selected D2D indicators

Process, structure and team characteristics summary

In summary, there is evidence of progress in measures of process and structure related to D2D (see Table 5-3). Participation in D2D is high. Comments from members suggest that simply being asked by AFHTO to participate made a difference. This suggests that simply asking members might be a useful strategy to encourage participation in other activities. Physicians respond to outreach from their peers. This suggests that physician champions could be important enablers of QI and further suggests that it might be useful to build physician engagement from existing relationships between physicians. QIDS Specialists were frequently mentioned in the D2D process and thus may be key enablers of QI. This suggests that providing QIDS Specialist-like help to teams without QIDS Specialists might advance measurement and improvement. There is an increasing frequency of conversations within teams over time. This is both a signal of success in meeting at least one of D2D's goals (that is, increasing conversations) and also validates the perception of D2D as an artifact. Next steps might include celebration of this success and taking even further advantage of the role of D2D as a conversation starter. There is evidence of improved team functioning, possibly related to increased conversations and more physician champions. This suggests that next steps focus on conversations and physician engagement. There is evidence of increasing data standardization in the absence of improved EMR functionality. This suggests an increased level of engagement in and effectiveness of people-based processes. Celebrating this may be useful to further encourage the spread and impact of people-based processes. Finally, the low response rate to surveys may have contributed to, and also resulted from, the fact that no stand-alone, overt reports on the surveys were produced. Next steps with understanding the impact of D2D need to examine impact of sharing results on participation in surveys. I also

am interested in examining my own decisions against producing survey result reports (see Chapter 9 for more reflexive examination of this).

Performance on primary care quality measures

In keeping with the focus of D2D to demonstrate the value of interdisciplinary teams in general (vs the value of a *particular* team), performance was considered at the membership level. This section describes average performance over all teams contributing data for EMR data quality, individual indicators, the composite quality score, per capita healthcare costs and finally, the relationship between quality and cost. Also presented is the reaction to the performance as reflected in the materials involved in communicating the release of each iteration.

EMR maturity

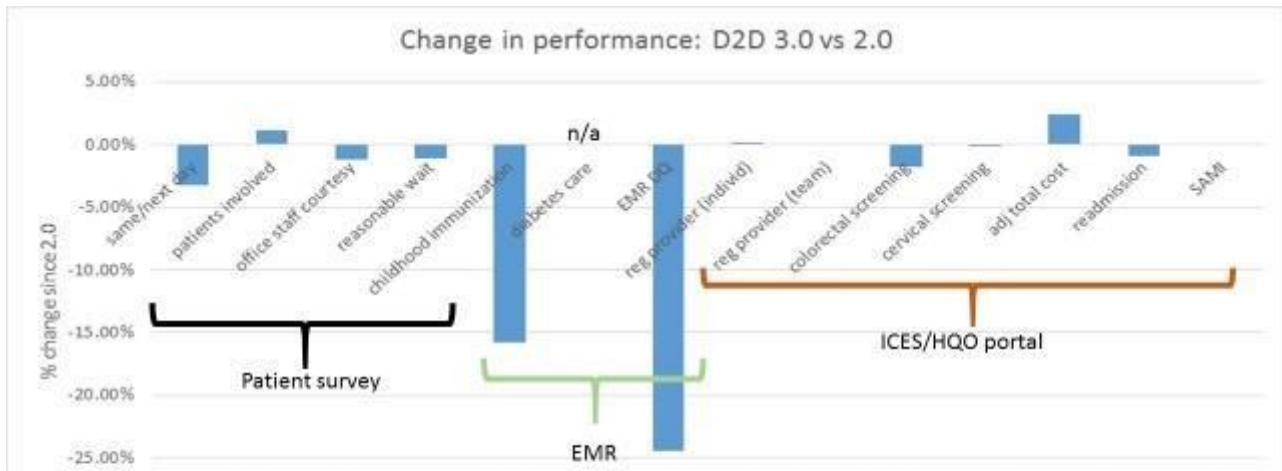
Contribution to the EMR data quality indicator was considered to be a surrogate measure of EMR maturity. The number of teams contributing to this indicator increased from 47 to 82 between 2.0 (when it was first introduced) and 3.0. This quantitative data quality score was presented for the team, the peer group and the entire D2D database in each iteration. Tracking the trend in EMR data quality over time was not meaningful because the components in this composite measure changed between iterations (see Figure 4-3). The EMR data quality measure appeared to have impact on teams. The launch of D2D 3.0 highlighted how one team found a systemic issue affecting flow of information from a lab to their EMR. They fixed the problem, improved their ability to detect colorectal cancer, and in their own words, “*saved lives!*”. The increased participation in measurement of EMR data quality and stories about the increased conversations about data quality over iterations of D2D suggest that *measuring* EMR data quality is having the desired effect of focusing the efforts of teams on *improving* data

quality. This is consistent with the concept of “enactment” (Weick, 1988), which was part of the foundation of the framework guiding this action research project. It therefore seems reasonable to continue explicitly measuring specific elements of EMR data quality and continually expanding the scope of the measure to ensuring continued meaningfulness.

Individual Indicators

In D2D 1.0, performance on the individual indicators in D2D was marginally better than comparator data. Performance in subsequent iterations of D2D did not change from the initial iteration (see Figure 5-6). Member reaction to performance varied across the first 3 iterations. In response to the wishful thinking of some members, I explicitly dispelled the hope that the reason for failure to show improvement was because teams were already performing well above target by showing that most indicators were below the targets for excellent performance set through a membership survey. Some noted that it might not actually be possible to improve performance in some areas because *“there are no evidence-based strategies for primary care to reduce readmissions or reduce cost. In addition, neither measure is stable at the practice level due to low volumes and high variation”*. Others felt it was *“early days”* and therefore too soon to expect improvements in performance. Although not shared externally, there was acknowledgement within AFHTO of the variation in performance between teams. The Board accepted the recommendation of the member-based steering committee that *“We need to address the variation between teams as a particular area for improvement”*. The Board concluded that *“the priority for the next iterations of D2D is to support members in quality improvement by focusing on reducing the variance in quality across teams”*. The lack of obvious improvement and the reaction to this confirms the premise of this research that measurement alone does not result in improvement. Although

this may seem self-evident, it is nonetheless important as many performance improvement initiatives (including D2D) focus predominantly on measurement with little overt support or attention to the use of the resulting performance data to improve outcomes. This reinforces the need for attention for this kind of support in subsequent iterations of D2D.



- Take home message:** Participation in measurement is increasing and having an impact. The lack of movement on current measures (eg patient survey and EMR data) suggests it is time now to expand the focus beyond measurement to improvement.
- More detailed comments:
 - Essentially same performance as 2.0. Most of the observed differences are <3% (which is within expected variation).
 - Exception 1: Childhood immunization. D2D 3.0 measures compliance with PHAC guidelines physician incentive bonus (as reflected in 2.0).
 - Exception 2: EMR Data quality. The 3.0 measure includes 2 additional components to expand the scope of the indicator beyond cervical cancer screening (which had the same high level of quality in 3.0 as in 2.0). See very positive impact of this change in story to follow.
 - ICES data is same time period as 2.0. Lack of difference in 3.0 performance suggests that the sample for 2.0 had about the same performance as the sample for 3.0, even with new teams added and thus is probably a reliable estimate of performance across the membership.

Figure 5-6: Change in performance over iterations

Quality composite measure

Many teams participated in the composite quality measure by contributing data to the “expanded” set of indicators and agreeing to be included in the calculations. Nevertheless, there was little membership reaction to the introduction of QRU. Questions from members first arose about the QRU around the time of *second* round of data submission for it (circa D2D 3.0). At that time, teams asked why and how the QRU was calculated and how they

could or should use the score. Comments from members indicated exasperation with and dismissal of the QRU:

“We don't use the roll up indicator. Haven't figured out how/why it's important and what we can do with it”.

“hard to believe a composite score does more good than bad”

The general approach to managing the concerns regarding the usefulness of composite measures at the front line was to side-step them and focus first on the value of the composite at an aggregate (membership) level. This is consistent with the “get started” philosophy that pervaded D2D. It is also consistent with the idea of embracing resistance rather than trying to overcome it, an idea that pervades the understanding of change as a continuous phenomenon (Thomas et al., 2011). It was possible to consider the composite quality measure at the aggregate level because the persistently high levels of participation in D2D had resulted in a dataset large enough to support reliability analyses. In contrast to the perception of the QRU at the team level, there was a higher level of interest and sense of encouragement in the comparison between AFHTO teams and the Ontario average. Some AFHTO Board members even suggested a public press release. Even still, this observation was shared tentatively. The caution was related to the different demographic characteristics of patients of teams relative to the population of the whole province. These differences raised the possibility that the observed differences in quality might be influenced by differences in demographics as much as by differences in care. The bottom line is that front line providers were not yet convinced of the usefulness of the composite measure. External observers were similarly ambivalent about the value of the composite measure in reflecting quality of primary care. The ambivalence may be related to the muted reaction of front-line providers, the demographic

differences among patients of AFHTO teams and/or the lack of data from other models of primary care. Further validation of the indicator itself, changes to increase its usefulness with front line providers and extension of the measure beyond AFHTO members might all be worth considering to increase the value of the composite measure as a reflection of quality in primary care.

Cost

The reaction to the decision to include cost ranged from “*how is cost relevant in a performance or quality measurement report?*” (remarkably enough from government representatives) to “*there is no way we can control costs – it's not fair to hold us accountable for that*”. Nevertheless, the Board of AFHTO agreed that it was important to include cost among the D2D indicators. The initial observation that average cost was slightly more for AFHTO members than for other models persisted across all 3 iterations. The cost data were met with the usual and predictable comments about data quality, including the “my patients are sicker than yours” rationalization of the higher costs for some teams. They were also met with disappointment on the part of the AFHTO Board, who felt the data conflicted with their belief in the value of team-based primary care. They did, however, recognize and express externally that the willingness of AFHTO to measure and report cost, no matter how disappointing the numbers, was part of the value and leadership of teams to the healthcare system. There was intense debate around the cost data, suggesting that including cost contributed to the achievement of D2D’s goal of generating conversation. Next steps might be wise to embrace the passion related to cost data as a mechanism to engage providers beyond AFHTO.

Quality and Cost

Although average costs in D2D were not lower than for the whole primary care sector, AFHTO members were able to demonstrate that higher-quality care (as demonstrated by high QRU scores) was associated with lower *per capita* healthcare costs (see Table 5-2). Multivariate linear regression modelling showed that patients of teams with higher quality scores had lower per capita healthcare costs, even when factors such as rurality, patient complexity, teaching status, team size and EMR maturity were considered (Mulder et al., 2016). This echoes the findings of Starfield (2009) who reported this relationship at national and regional levels. The reaction to the results of the QRU analysis predictably centered on concerns about the quality of the data, which were mostly couched in questions about how the indicator was calculated (which was admittedly complex -- see Chapter 4). Even after these concerns were addressed through posting responses to Frequently Asked Questions and a series of videos describing the QRU (AFHTO, 2016), the findings did not immediately get high profile in AFHTO communication. For example, the only mention of these findings at the 2015 AFHTO conference consisted of references to 2 slides in the middle of one 15-minute address to the plenary session. In contrast, all 9 of the abstracts I prepared that focussed on these findings were accepted for presentation at premier Canadian and North American primary care conferences (see Appendix 4). One was awarded “distinguished lecture” status as one of the top 4 papers submitted (CFPC, 2016). It was only after these presentations that members began asking AFHTO staff to highlight the quality-cost observations in a more high-profile way to members and Ontario partners and stakeholders. Something seems to prevent AFHTO members from taking their own success in

demonstrating value seriously. This has implications for persistent credibility with external stakeholders and thus must be addressed in future efforts with D2D.

Table 5-2: Regression of quality on per capital health care cost (Mulder et al., 2016)

Dependent	r ²	coefficients (significant at p<0.05)			r ²	Coefficients (significant at p<0.05)			
		Quality	SAMI	Hospital-EMR link		Quality	SAMI	Hospital-EMR link	Patients served
		Urban (n= 49)				Rural (n= 47)			
Total costs without institutional	0.463	-0.454	0.432		0.613	-0.204	0.696	0.233	
Services costs	0.424	-0.434	0.421		0.485		0.697		
Primary care costs	0.005	0.067 (p>0.10)			0.443	0.024 (p>0.10)	0.340		0.531

Performance measures summary

In summary, performance is slightly better than comparators on some, but not all indicators. AFHTO members experienced the same level of performance on individual primary care outcome measures throughout the first 3 iterations of D2D. In addition, variation between teams is considerable. This illustrates that measurement alone does not result in improvement and further, that there is a need to support improvement efforts in addition to the

measurement of performance. There was no change in performance on the composite quality measure either. However, there was evidence that higher quality was related to lower *per capita* healthcare cost. Nevertheless, sharing of information about the quality composite measure and its relationship to cost was tentative. This suggests that AFHTO members and external observers might not be convinced of the usefulness of the composite measure.

Validating, increasing the usefulness of composite indicator, and extending it beyond AFHTO members might increase its perceived value. Nevertheless, there is evidence in D2D of progress in process measures. For example, there is increased participation in the measurement of EMR data quality and a concomitant increase in interest in improving it. This suggests that it is worthwhile to continue expanding the definition and measurement of EMR data quality. There was intense debate about the cost indicators, suggesting that including those indicators generated the expected increase in perception of cost as a part of quality. Future iterations could embrace the passion in the cost conversations to expand the view of quality in the sector to include consideration of cost. Finally, the data show that the AFHTO does not hold D2D as its highest priority, nor does it leverage it as extensively as it could for advocacy, despite naming measurement as a strategic priority, and despite the success of many aspects of D2D. This could mean that AFHTO members have difficulty accepting evidence of their own success. It certainly suggests that next steps with D2D need to address the perceived priority of D2D among AFHTO members and staff. Table 5-3 summarizes the observations regarding measures of process, structure and performance (i.e. outcomes) along with high-level reflections and implications for the next action cycle in this action research project. The next section describes evidence from conversations that helps to identify enablers for QI.

Table 5-3: Summary of observations, reflections and implications of process and structure measures

Measure	observation	reflection	Implication
Process, structure, team characteristics			
Response rate for team characteristics surveys	Response rate was persistently low. No standalone reports produced.	Lack of reports may have contributed to low response rate; counter to “getting started”	examine impact of sharing results on participation
Participation	Participation is high; “AFHTO asked me” is key rationale	asking encourages participation –who asks who might matter	leverage AFHTO’s willingness and ability to ask teams to participate.
Physician champions	Physicians respond to outreach from their peers.	physician champions could be important enablers of QI.	build physician engagement from existing relationships between physicians
QIDS Specialists	QIDS Specialists frequently mentioned in the D2D process	QIDS Specialists may be enablers of QI	Provide QIDS Specialists-like help to teams without QIDS Specialists.

Team climate	Improving scores over time	related to increased conversations and more physician champions.	focus on conversations and physician engagement
Conversations	Increasing frequency of conversations over time	success in achieving key D2D goal; validates role of D2D as an artifact that stimulates engagement.	Celebrate! to take advantage of D2D as a conversation starter.
Access to EMR data	Increasing data standardization without	increased engagement in and effectiveness of people-based processes	Celebrating to further encourage these people-based processes.
	improved EMR functionality		
Performance			
EMR maturity	increased interest and participation in measuring EMR data quality	measuring EMR data quality is increasing interest in data quality	continue measuring and expanding EMR data quality
Individual D2D indicators	lack of obvious improvement	measurement alone does not result in improvement	Support improvement in addition to measurement of performance

Quality	Sharing of information about quality composite measure was tentative	Front line providers and external observers not convinced of the usefulness of the composite measure.	Validate and increase the usefulness of composite indicator; Extend it beyond AFHTO members
Cost	intense debate around the cost data	including cost in D2D generated the expected attention to cost as a part of quality	Embrace the passion in the cost conversations to expand inclusion of cost in measuring quality
Quality and Cost	organization does not prioritize D2D or leverage it for advocacy	AFHTO members have difficulty taking their success in demonstrating value seriously.	Increase perceived priority of D2D among AFHTO members and staff

Conversations

Overview of data extraction process

Based on the selection process described in Chapter 3, 842 emails were selected and reviewed over the 3 periods (D2D 1.0 – 3.0). The first level codes in the initial coding template were based on the research framework described in Chapter 3. A second level of codes were identified on review of the data. These codes were subsequently merged into aggregate dimensions, according to the process described in Chapter 3. The four aggregate dimensions which emerged from this process were building relationships, help-seeking and self-reliance, diversity of perceived priority of D2D and the importance of getting started with

QI activities. The final version of the coding template with representative text for each Level 2 code is shown in Appendix 2. Elements of these themes in the quantitative data, where apparent, are presented below along with the evidence regarding the themes emerging from conversations. This is because making meaning of the whole experience with D2D requires consideration of all the data, quantitative *and* qualitative. The quantitative data referred to below does not include performance data because there was no evidence in that data source to support the themes emerging from the conversations and the process measures. This underlines the importance of looking at more than quantitative performance to understand the experience of primary care providers in measuring and improving performance.

Aggregate dimensions/master themes

Building relationships

One of the themes emerging from the data was that of “building relationships”. There was evidence of informality in exchanges between me and members and external stakeholders. Exchanges were characterized by self-deprecation (“*this might be a silly question but...*”), excitement (“*this could be a lot of fun!*”) and agreeable enthusiasm (“*Sure thing!*”). There was overt reference to familiarity between me and respondents in the exchanges: “*Of course! I always do what Carol asks me to*”. Early in the D2D trajectory, these light-hearted agreeable exchanges were often requests for, or responses to, introductions, especially with respect to clinicians or others from whom input was needed. Over time, conversations evolved to include attempts to take advantage of relationships (“*Hi friends. I hope you can help me send this out to your members.*”). Less frequently, the exchanges were responses of AFHTO staff to requests for help in facilitating discussion or resolving conflicts. AFHTO members were frequently asked for their input by AFHTO staff via surveys and

individual or group conversations via email, phone or face-to-face meetings. AFHTO members also spontaneously provided input through similar channels.

There is evidence of a high degree of trust among members that AFHTO staff would not only hear, but be directed by input from members, as demonstrated by one member who said

“...after much discussion, it was decided to remove [an indicator] from D2D. I have to say that I was delighted with the process as we are trying to keep D2D “meaningful” and the providers' voices were listened to”.

The nature of the feedback was occasionally critical of decisions made by AFHTO and external stakeholders alike. Members cited distrust of external bodies as reasons for not participating in D2D or other AFHTO activities. This is clearly shown in a warning from one member: *“Make no mistake, AFHTO is an arm of the government!”*. The expression of these and similarly negative sentiments illustrates in a perverse way the importance of trust and good relationships in achieving engagement and participation. Even arguments and debates were perceived to be useful in building relationship. One member notes that it was *“interesting that anyone could be this belligerent about measurement. I didn't think that anyone really cared that much, so I take this as a good sign”*, illustrating a (possibly unconscious) appreciation of the value of ambivalence and embracing argument in change as an ongoing phenomenon. Comments like *“AFHTO is very credible so I will do it for them”* suggest that the relationship of members with AFHTO staff was a key driver in their decision to participate in D2D. Quantitative data that aligned to these ideas included the growing and sustained participation in D2D, the increasing frequency of conversations about measurement within teams, improving team functioning and the impact of a warm hand-off on physician engagement.

The focus on relationships is consistent with AFHTO's mandate as a membership association. The evolution from starting relationships (via introductions) to leveraging them to get help from others implies success in building relationships. Next steps with D2D can, and should, build on the strength of the relationships between AFHTO staff and members.

Help-seeking and self-reliance

Another theme that emerged was that of a desire for help in solving problems. While the theme initially emerged as pleas for help from AFHTO members, review of the entire data set showed that this theme applied to AFHTO staff and other stakeholders as well. Evidence of help-seeking behaviour was ubiquitous, ranging from "*Who is responsible for covering for the cost to the EMR vendor, the host FHT?*" to "*All car rentals in [northern town] limit the number of km driven in a single day to 100 km, leaving me to pay 20 cents/additional km. Not clear if you are or are not paying to fill the gas tank*" to "*I am logged in as a member and unable to launch the tool...*". There were also abstract references to what were perceived to be insurmountable problems: "*I'm worried about the cost*" or "*I can't find any indication that [my board members] were ok for me to send it to you so I can't send it*". These were not accompanied by requests for help, suggesting a desire to just share the problem even if it could not be solved.

Another phenomenon was the overt request of QIDS Specialists for help with what appeared to be technical issues, but were in fact requests for emotional support. Much of the help-seeking behaviour was misdirected. Members sought help from AFHTO for issues that were very intentionally out of scope for AFHTO staff to address. For example, QIDS Specialists were hired directly by the teams they served to ensure autonomy of AFHTO members in recruiting and managing their own staff. Nevertheless, members directed

questions regarding operationalization of QIDS Specialists roles to AFHTO staff instead of addressing them locally as they had every right and responsibility to do. Another demonstration of ineffective help-seeking behaviour was the expressed interest in hearing “success stories from peers” and the spectacular failure of multiple efforts to persuade members to share their stories. A third example of unproductive help-seeking is the persistent preference for email for AFHTO communication and the equally persistent low readership rate (approximately 40% were opened).

There was evidence that the emotional support and encouragement offered (without technical advice on the actual problem) was appreciated and effective in promoting local action on the problem. There were many comments like “*Thanks for coming up to visit us. It was great to be able to spend a bit of time chatting with you*”. Over the trajectory of D2D, problem-solving behaviour changed to include sharing of questions (“*Does anyone have a way of tracking smoking status in [a common EMR]?*”) and solutions (“*I can share the process map (swim lane diagram) later this week if you are interested*”). This was especially obvious among QIDS Specialists and related to issues with accessing data. The quantitative data were silent on the help-seeking behaviour. They did, however, provide hints of self-reliance in the form of increasing data standardization (a people-based process) in absence of technical solutions for limited EMR functionality.

The tendency of asking for help instead of trying to find solutions (especially initially) may be a signal of lack of readiness for change in addition to actual gaps in capacity regarding the problems. Since solutions to the problems might not actually address the readiness for change and vice versa, it is important to examine the evolving nature of “help seeking” in subsequent iterations of D2D.

Diversity in perceived priority of D2D

A third theme emerging from the text is the diversity of perceived priority of D2D among AFHTO staff, members and external stakeholders. Conversations suggest that while D2D was important operationally, it was not universally seen as a strategic priority. In early iterations, most of the conversation about D2D was internal to AFHTO staff and much of that initiated by me, as the QIDS program lead. The dialogue was about the mechanics like “*What’s the date for D2D indicators announcement?*” or “*I know [you] have been crunching with the D2D materials today. Just wondering if you’re still aiming to get the slides to me tonight, or should I look for them in the morning?*”, not the strategic aspects of D2D. In contrast to the high volume of these operational discussions, there were few mentions in internal AFHTO conversations regarding the role of D2D as a strategic priority.

Conversations about D2D with members were also mostly about the mechanical aspects of the initiative such as those reflected in these comments from members: “*unfortunately even with the extension I cannot submit data by July 25*” and “*Are we still allowed to pick and choose (as we did in 2.0) what we want to report?*”. A new aspect of conversations about D2D that emerged after the first iteration and persisted through subsequent iterations was technical problems with the functionality of the D2D reporting tools (data submission and presentation). Emails to members about D2D that were not about technical problems were occasionally deferred as shown in comments like “*Sorry, but [other topics] are the priority*”.

Initially, external communication about D2D was scant. Over subsequent iterations, external messages about D2D were more comprehensive, dedicated only to D2D and sent to extensive distribution lists. Conversations initiated by external organizations centered

initially on AFHTO's ability to provide credible input from primary care providers.

Examples include invitations to represent front-line providers on provincial committees and multiple requests for AFHTO to sign off as a “knowledge user” to meet the requirements of research funding agencies. External organizations were willing and able to provide help to AFHTO, but it was primarily service-oriented and directed towards operational or technical issues: “*we have the ability to export the data you are looking for*”. There was gratitude for the provision of input, but D2D was virtually never raised by external stakeholders over the first 3 iterations of D2D. One partner conceded that “*There are no current plans [to follow AFHTO’s lead]. We could raise it for discussion but it won’t happen soon*”. Annotation of AFHTO’s activities was added to the graph showing uptake of the Primary Care Practice Report only after explicit request to the organization producing it, even though the graph already included annotations regarding activities of others (see Figure 5-2). AFHTO seemed to be considered as a useful partner, but not a leader.

In contrast, conversation about D2D by members suggested they saw D2D as a valuable strategic priority. One member noted that “*The beauty of D2D is the snapshot of provincial primary care it generates, the gap analysis it allows (provincially and in individual FHTs) and the opportunity to address provincial primary care gaps in a structured QI way.*” Members felt that “[D2D] may put AFHTO in a strategically important place, so nice work!”. They felt D2D was a tool AFHTO could use to advocate on their behalf.

The internal and external messaging about the vital strategic importance of measurement was not always manifested in AFHTO staff priority-setting nor in the perceptions of D2D by external stakeholders. This phenomenon was not unique to D2D. There was a similar paradox in the expressed vision for patient-centered care and the scarce

evidence of direct patient engagement by AFHTO members². This range in perceptions was mirrored in the limited external distribution for initial iterations of D2D and the growing and sustained participation over time. These observations suggest a need to address the gap between the vision of AFHTO and decision-making among AFHTO staff and members regarding the strategic priority of D2D and patient engagement.

The importance of getting started in QI activities

The fourth theme relates to progress with QI activities and the value of getting started. Quantitative participation data illustrates that there was progress in QI activities such as accessing data, talking about performance and contributing to measurement (i.e. via D2D). Qualitative data showed that members perceived value in measurement and in particular, their ability to “*see where they stack up*” with peers, implying an interest in performing well or taking steps to achieve that. There was a distinct focus on the value of getting started, even in the face of incomplete information or engagement. Initially, this was most obvious in messages from me, possibly due the higher overall volume of communication about D2D from me in my role as program lead relative to other AFHTO staff or members. D2D 1.0 was formally framed as “*a way to get started*”. The concept also showed up in comments of members such as: “*I think messiness may be worth trying*” or “*I know it is not perfect, but we are moving miles ahead*”. There was evidence that getting something underway worked to generate enthusiasm. One member reflected that

² Patient engagement in this context refers to partnership and shared leadership with patients and similar higher order activities (relative to consultation or involvement) on the continuum of patient engagement (Carman, et al., 2013). Consultation and involvement are not uncommon among AFHTO members individually and collectively through patient surveys and incorporation of patient feedback into processes such as the composite quality measure.

“a few months ago most of my ED’s are like. “D2D? What’s that? More work? yuck....!”

Now they are all like “hey can we submit to D2D?” after seeing it”.

Nevertheless, there was minimal evidence of celebration of progress with D2D. There were suggestions about celebrating or thanking various players, but little actual activity or investment in the idea. For example, the launch of D2D 1.0 proceeded with no other celebration than a hand-written post-it note from a colleague. While the launches of 2.0 and 3.0 received slightly more fanfare through cupcakes made by a QIDS program staff member (see Figure 5-7) there was no formal ceremony or recognition of these milestones from AFHTO leadership or Board.



Figure 5-7: Homemade cupcakes provided by AFHTO staff member to celebrate launch of D2D 2.0

The growing participation in QI activities was not accompanied by an expanded scope of conversation about QI beliefs. For example, questions about the new composite quality indicator (a novel and complex concept) focussed on the mechanical aspects rather than the

strategic value or theoretical rationale. An example was “*What does ‘Please check here if you would like your data to be included in the quality roll-up indicator’ checkbox means?*”. Neither was there evidence of moving beyond measurement to using data to improve clinical performance. One team reported that “*patient experience surveys were distributed, but when collected in they were placed in a box in [the EDs] office, they weren't analyzed*”. This was mirrored in the static performance of D2D indicators over the first 3 iterations. Members provided multiple reasons why performance had not improved yet and might not in future iterations. The governance committee for D2D felt that it was unreasonable to expect performance to improve until at least 2 more years had passed. The fact that the obvious ability to “get started” with measurement did not translate into improvement may be related to the early stage of the D2D initiative (just 3 iterations in only 18 months at time of writing). The early stage in the life cycle may also be the explanation for the relative absence of interventions or enablers for improvement.

Summary of conversations

In summary, conversations over the first three iterations highlighted four themes. There was evidence of a focus on relationships. This is consistent with AFHTO’s mandate as a membership association. The observed evolution from starting relationships (via introductions) to leveraging them to get help from others implies success in building relationships. It illustrates the value and impact of facilitative and nurturing role of AFHTO staff, another key feature of change leadership when viewing change as always and already happenings (Tsoukas & Chia, 2002). Next steps with D2D can build on the strength of these successes with relationships, especially those between AFHTO staff and members.

There was a tendency for staff and members to ask for help instead of trying to find solutions, especially early in the D2D initiative. This may be a signal of lack of readiness for change in addition to or instead of actual gaps in capacity regarding the problems. It also points to the value of nurturing in supporting change, as suggested by (Palmer & Dunford, 2008) especially since there was evidence of emerging self-reliance, especially among QIDS Specialists who were actively facilitated by AFHTO staff. Next steps should examine the evolving nature and rationale of “help seeking” and self-reliance behaviour.

The conversations showed that D2D was not the only or first priority for AFHTO staff or external stakeholders, even though many members treated it as a high priority within their teams. Perhaps D2D was not sufficiently disruptive outside of AFHTO to serve as effectively as an artifact as it did internally. There was also low interest and participation in patient-engagement among members. These observations seem to be a mismatch with AFHTO’s strategic priorities for measurement and patient engagement. Addressing the apparent gap between the vision and on-the-ground decision-making regarding D2D and patient engagement might be worth pursuing in subsequent iterations of D2D.

There was demonstrated intent and ability to “get started” with measurement. This did not translate into improvement or even availability of supports for improvement. This phenomenon may be related to the early stage of the D2D initiative which, at time of analysis, was on its 3rd iterations and into its 18th month. Next steps should leverage the success of the “get started” approach and apply it to specific activities to support improvement.

Table 5-4 summarizes the observations emerging from the conversations along with relevant quantitative data consistent with the themes emerging from the qualitative data, the high-level reflections and implications for the next action cycle in this action research project.

Table 5-4: Summary of observations, reflections and Implications of conversations

Theme	Data from conversations	Data from participation and process measures	observation	reflection	implication
Building relationships	Increasingly productive relationships and evidence of negative impact of bad relationships	Growing and sustained participation and increasing relationship-based behaviours (conversation, team-functioning, physician engagement)	focus on relationships with evolution from starting relationships (via introductions) to leveraging them to get help from others	consistent with AFHTO's mandate as a membership association, evolution implies success in building relationships.	build on the strength of the relationships between AFHTO staff and members.
Help-seeking and self-reliance	Misdirected requests for help contrasting with local self-initiated problem-solving	No evidence related to help-seeking, but hints about self-reliance through increasing data standardization in absence of improved EMR functionality	tendency of asking for help instead of trying to find solutions (especially initially)	may be a signal of lack of readiness for change and/or actual gaps in capacity regarding the problems.	examine the evolving nature of “help seeking” needs in subsequent

					iterations of D2D.
Diversity in perceived priority of D2D	D2D not always the only/highest priority internally, but perceived as strategically valuable by members	Limited initial external distribution contrasting with growing/sustained participation with increased data submission and increasing frequency of conversations	D2D not highest priority for AFHTO staff or external stakeholders even though members see it as a priority; low interest in patient-engagement among members	Mismatch between AFHTO's strategic priorities (measurement and patient engagement) and actual activities of AFHTO staff and members	address the gap between the vision and decision-making regarding D2D and patient engagement.
importance of getting	Interest in getting started with a good	Increasing participation in many aspects of D2D	demonstrated ability to “get started” with	may be related to the early stage of	leverage the success in

started in QI activities	attempt rather than waiting for a perfect solution and appreciation for evidence of small improvements		measurement did not translate into improvement or supports for improvement	the D2D initiative (3 iterations, 18 months)	“getting started” and apply it to specific activities to support improvement.
--------------------------	--	--	--	--	---

Phase 2 summary

Review of the quantitative and qualitative data about the experience with the first three iterations with D2D generated a series of observations. These observations were consolidated into themes or aggregate dimensions. As shown in Table 5-4, these dimensions were building relationships, help-seeking and self-reliance, diversity of perceived priority of D2D and the importance of getting started with QI activities. Critical reflection on each of the individual observations suggested implications for each (see Table 5-4). In Chapter 6, these observations and their respective implications are considered together to generate meaningful suggestions for action.

Chapter 6 Phase 3: Preparation for 4th iteration

Preamble

Chapter 6 focuses on Phase 3 of the action research. Phase 3 involves a deeper reflection on the experience of AFHTO with the initial iterations of D2D, described in Chapter 5, and describes actions emerging out of these reflections for consideration in Phase 4. The experience with the initial iterations of D2D (Phase 2 of the action research project) highlighted themes to guide the focus of further actions. These include: building relationships, helplessness and self-reliance, diversity of perceptions of priority and the importance of getting started with QI activities. This descriptive information is essentially an answer to “What happened?”, the first of three questions in the critical reflection framework defined by Rolfe et al. (2001). Chapter 5 also introduced reflection on the observed experiences and their implications for further action, addressing in part the second and third questions in the framework: “So What?” and “Now What?”. These reflections were mostly at the level of individual observations. In Phase 3 (that is, this chapter), these observations are considered collectively and in the context of evidence from the literature.

While many of the observations were consistent with, and therefore reference the literature reviewed at the outset of this research, several of the themes emerging from the data were not anticipated and prompted review of additional literature, which is discussed in the context of the reflections below. This deeper reflection supports attempts to make meaning of the experience with D2D in a more comprehensive way. These reflections refine the themes emerging from Phase 2 to the point of suggesting specific actions for future iterations of D2D.

As noted in the methodology discussion (Chapter 3), the actions were suggested on the understanding that their value would be reflected by the extent to which they were considered and/or implemented in subsequent iterations.

Reflections and implications for next steps: Relationships

The emergence of relationships as important in D2D is consistent with the desire of psychological safety in the context of change (Schein, 1999) and importance of nurturing in the process of continuous change as described by Palmer & Dunford (2008) and Weick & Quinn (1999). The evidence of members' trust of AFHTO staff suggests that the goal of engaging with members and being (or at least appearing to be) responsive to member direction has been met. Improving team climate may be related to better relationships with physician champions and increased conversations within teams about performance. The focus of AFHTO staff on relationships is consistent with the role of secretariat staff working on behalf of members, which is appropriate given that over half of the budget for AFHTO staff comes from membership dues. Members are not required to join the association, nor participate in, or otherwise support, activities of/requests from AFHTO staff. Similarly, external stakeholders are not required to collaborate, or even communicate, with AFHTO members or staff. Progress with collective action therefore depends on success in engaging others voluntarily. Social capital theory suggests that relationships are the preferred currency for success in initiatives that cannot be successfully implemented by a single organization (Lesser, 2000). The observed evolution of relationships from introductions to the making (and acceptance of) requests is a signal of success in building relationships (Misner, 2011). The age-old adage that "*who you know matters as much as what you know*" may explain the responsiveness of AFHTO members to requests from

AFHTO staff. This is consistent with literature showing that conversations among colleagues is an important driver for change among physicians (Ivers et al., 2014). It seems that it is important to ask people to do things. It also seems like the person who is asking matters as well. The easy-going, light-hearted tone of the conversations is also consistent with the necessary dependence on voluntary engagement in AFHTO's work. As noted above, AFHTO staff are not able to make demands on their employers, funders or stakeholders. The data show that AFHTO staff use conversations and, in particular, fun conversations to engage stakeholders. Owler et al. (2010) (who reviewed nearly 100 papers over the last 20 years) and Tews et al. (2014) are among authors describing how fun improves job satisfaction and energy and reduces anxiety and burnout. They also distinguish between "organic" fun, which emerges spontaneously and is universally appreciated and externally defined planned fun like barbeques or celebrations, that are more likely to be mocked than enjoyed. They also caution against considering fun as a panacea for all workplace problems. Baptiste (2009) specifically recommends against prescribing "silly hat days" at the expense of being attentive to other material needs of staff. In keeping with the idea that not all "fun" is equally effective, not all conversations support change. To be effective as tools for change, conversations need to rise beyond mere chatter. Conversations can be important vehicles for change (Capelli & Smithies, 2009; Macpherson et al., 2006), but they need to be 'conversations-for-action' (Dervitsiotis 2002, p. 1088) or "conversations for performance", in the words of Ford & Ford (1995, p. 549). This requires advanced listening skills and attention to new possibilities (Palmer & Dunford, 2008).

Relationships were important to providers too. Many providers espouse the principles of Barbara Starfield that the patient-provider relationship is the core value of interdisciplinary primary care (Premji & Hogg, 2016). The importance of physicians' relationship

s is well described in the literature with physician-to-physician communication identified as a key enabler for the audit and feedback (Ivers et al., 2014) and academic detailing (Allen et al., 2007) approaches to improving outcomes. It is therefore not surprising that the provider focus on relationships extends beyond their patients to their interactions within and between teams. The experience with and evidence about relationships had implications for next steps with D2D. These include the following:

- Nurture relationships: In keeping with the recognized value of nurturing leadership when supporting change as a continuous phenomenon (Palmer & Dunford, 2008), continue investment in building and, especially, sustaining relationships. Be more attentive to and responsive to elements of fun to deepen relationships and thus contribute to resilience and interest in innovation.
- Maintain attention on conversations: Continue to support what Dervitsiotis (2002, p. 1088) refers to as “conversations for action”. Paying more attention to conversations among people and on topics where more action is needed might help increase the ability of the conversations to lead to the desired changes in measurement and performance.
- Continue inviting providers to participate: Keep inviting members to participate, even with the existing high levels of participation and even on topics that are not fully understood or even necessarily embraced by the individuals or organizations being requested to act. To increase the chance of positive outcomes, pay attention to who is asking and who is being asked.
- Build cadre of physician champions: Increase opportunities for physicians to reach out to peers they know to initiate conversations and otherwise encourage participation in measurement and improvement activities.

Reflections and implications for next steps: help-seeking and self-reliance

The emergence of help-seeking behaviour as an influence on D2D is not surprising given the understanding of change as a function of the balance between the perceived need to change (survival anxiety) and the anxiety associated with learning new things associated with that change (Smith & Elliott, 2007). The fact that much of this behaviour was misdirected or otherwise dysfunctional highlights the potential that the help-seeking behaviour was related to what Armenakis & Bedeian (1999) described as the “psychological pain” associated with change. The pattern of help-seeking behaviour suggests that *expressing the request* for help is more important or valuable to members than *receiving the support* requested. This pattern of behaviour is recognized in change management literature. For example, using politically safe reasons to defer change (that is, citing workload as a barrier or asking for help) is consistent with the unreasonably reasonable behaviour that Ezzamel et al., (2001) describe as a form of resistance to change. The tendency to focus on problems without hope or expectation of resolution is emblematic of what Prochaska & DiClemente (1992) describe as a “resigned precontemplation”. This is the earliest stage of self-change in their theory of change and is characterized by belief that change is either not necessary or, in this case, not possible. It is therefore not surprising to see this pattern in the earliest days of the D2D initiative. While these two perspectives (Ezzamel et al., 2001 and Prochaska& DiClemente, 1992) differ in important ways, they converge on the conclusion that responding to specific, overt requests for help will not likely support change and may in fact, perversely act against the change.

The observed gratitude for emotional support (without accompanying technical help) was a hint that some of the help-seeking behaviour might be as much about a lack of confidence as a lack of competence. The evolution of requests for help to requests for confirmation of proposed

solutions suggests that members knew what to do (in that they were able to propose solutions), but wanted assurance or external validation. The role of confidence in behaviour change is particularly prominent in the “contemplation” stage of change in Prochaska & DiClemente’s model (1992), which is characterized by low confidence in one’s own ability to make desired changes. Confidence in one’s ability to undertake a new behaviour is identified as an important enabler of change in many other theories of change such as the health belief model (Rosenstock, 1974), the theory of planned behaviour (Ajzen, 1991) and the health action process approach (Schwarzer, 2008).

The self-reliance that appeared to emerge over time among QIDS Specialists (for example, sharing solutions, not just questions and active efforts to change) is more characteristic of people in later phases of readiness to change. These equate to the “preparation” and “action” stages identified in Prochaska & DiClemente’s model of change (1992) and “intenders” or “actors” in the Health Action Process Approach (Schwarzer, 2008). Both theories hold that people in these stages have higher levels of confidence and receptivity to encouragement and reassurance. The higher self-reliance among QIDS Specialists who were starting to solve their own problems is consistent with the increased control and power that people feel over their environment simply by taking action, according to Weick (1988). The intense support provided to build the community of practice among QIDS Specialists through weekly meetings, field visits, knowledge transfer and exchange sessions and other direct engagement was therefore an appropriate, and apparently effective, strategy to support change among QIDS Specialists. When considered as a spectrum ranging from help-seeking to self-reliant behaviour, the experience of members seems aligned with theories of behaviour change such as those described by Prochaska & DiClemente (1992), Schwarzer (2008) and more historically, Rosenstock

(1974). These theories all recognize that people go through different stages in the process of behaviour change. More importantly, they also recommend stage-specific strategies to support individuals in changing their behaviour. For example, the literature suggests that effective processes to support change among people early in the change process (e.g. contemplation or pre-contemplation stages as defined by Prochaska & DiClemente, 1992) are those focussed on helping people re-evaluate their social environment peer pressure or consider risks. The literature suggests that providing them with solutions (the historical and default response of AFHTO staff based on a desire to serve their members) not only does not help but might deter desired behaviour change. Matching support strategies to readiness to change is the most effective way to support change, even in the face of what appear to be direct requests for help. Help-seeking may also be an example of what Ezzamel et al. (2001) describe as reasonably unreasonable behaviour. For example, the claim that teams “couldn’t do [D2D] without QIDS Specialists” is not consistent with the relatively large number of teams without QIDS Specialists who do in fact contribute data to D2D. This raises the possibility that asking for help (which signals willingness but inability to participate) might be more reasonable than overtly declining to participate. If the underlying issue is not the reasonable need for help, but the less politically safe lack of interest in participating, providing help might have unintended consequences. For example, members whose coping strategy is to ask for help might paradoxically resent the provision of help which essentially destroys their coping strategy. Providing help might therefore tip the balance of forces at play in the ambivalence about D2D away from participation and paradoxically interfere with, rather than support, the change associated with D2D. This is doubly ironic as AFHTO staff are driven to provide help as part of their commitment to good customer service to their members.

By virtue of choosing to ask questions rather than act, help-seeking behaviour might also be evidence of what Anderson (2003) called “decision avoidance”. Anderson (2003) and Mellahi & Wilkinson (2010) suggests that it seems easier for people to forgive themselves and others for bad results related to “sins of omission” (that is, decisions avoided or not made) compared to the same results related to overt decisions or “sins of commission”. Responding to requests for help from a person in this situation might end up forcing them into a decision they really don’t want to make, notwithstanding their request for that help. They might not appreciate that.

The above reflections have implications for next steps with D2D. They suggest the following actions:

- Match interventions to readiness to change: Design interventions to support members in participating in measurement and improvement according to the stage of change they are in. Consider matching interventions to the goal achievement orientation of teams. Manage the implications of this approach to supporting members for the usual customer service orientation of all staff in this member-owned association.
- Explore ambivalence: Recognizing the importance of embracing ambivalence in supporting change as a continuous phenomenon (Piderit, 2000), explore help-seeking behaviour to understand its potential role as a signal of ambivalence about measurement and improvement.

Reflections and implications: diversity of perceived priority of D2D

The high interest in D2D among distant external organizations suggests that D2D has achieved at least part of the goal of demonstrating the value of AFHTO. There was also evidence that AFHTO members prioritized D2D to the point of doing the work to contribute data

and reflecting on their performance data relative to their peers. This suggests that D2D was successful in achieving its goal of engagement and participation of members. This does not, however, suggest that members were convinced of the strategic value of D2D. There was ample evidence that participation was more a sign of compliance with perceived expectations of AFHTO staff. The high participation in the quality roll-up indicator processes well before members fully understood, or even could ask questions about how the indicator worked, is another hint that participation was about complying with expectations as much as engagement with the strategic purpose of D2D. The muted interest in the evidence of the value of teams that was emerging from D2D might be another hint that members were not acting out of strong engagement with the strategic role of D2D.

AFHTO internal staff had a more operational view of D2D than members did. This might be expected and appropriate for the stage the project was at, that is, introduction and refinement based on collaboration with early adopters. The perceptions of D2D among external stakeholders are more or less in line with the amount and nature of conversation about D2D with them by AFHTO staff, which is to say low level of awareness and framing as an operational vs a strategic priority. This is opposite from the apparent priority for measurement expressed in AFHTO's strategic priorities. This is not the only mismatch between what Argyris & Schon (1974) describe as espoused theory and theory in use. AFHTO's Board actively embraces Starfield's definition of the quality of primary care as being dependent on the quality of the relationship between patients and providers over time. The lack of traction of efforts to encourage patient engagement appear inconsistent with the express focus on patient-centered care in AFHTO's mission and strategic directions.

Argyris & Schon (1974) have observed that individuals and organizations are governed by two theories: the set of principles they *believe* they operate on (espoused theory) and the principles that guide their *actual actions*, consciously or otherwise (theories in use). He argues that the key to improved effectiveness is to eliminate the incongruence between the two theories. This is difficult and potentially embarrassing, particularly if the theory in use is not politically appropriate. Since avoiding embarrassment is a primary driver of organizational behaviour (Argyris, 1996), there is a tendency for people to engage in defensive routines to delude even themselves about what is really guiding their actions. A commonly used delusional tool is the affinity for rationality, in which uncomfortable or potentially embarrassing questions can be labelled as “irrational” or “emotional” and therefore dismissed (Vince & Broussine, 1996). For example, it is irrational to suppose that primary care providers do NOT want to demonstrate their value. By virtue of being irrational, this assumption can escape examination in the quest to understand theories in use and thus improve effectiveness.

Another phenomenon that can paradoxically keep individuals and organization from aligning their behaviour with their stated beliefs is behaviour similar to what Bazerman & Samuelson (1983) refer to the “winner’s curse”. Individuals or organizations may only realize that the costs of achieving a goal were too high once they have achieved it. Among the possible reasons for this is ‘bounded awareness’ (Chugh & Bazerman, 2007) in which decision-makers miss readily available information, especially when they believe there is little risk that they have done so. As a result, winners (in whatever effort they have undertaken) may prevent themselves from realizing the benefits of having achieved their goal.

The observations and reflections regarding perceived priority of D2D suggest specific actions for next steps with D2D. These include:

- Explore the theories in use regarding measurement and improvement: There could be a mismatch between the espoused theory of AFHTO staff and their theories in use. Alternatively, not all staff and members may actually have the same espoused theory. Understanding differences in perceptions of priority therefore necessitates a deeper understanding of the beliefs that all parties espouse and what they actually act on in their daily lives.
- Communicate the value of D2D more effectively: Attempt to resolve the “winner’s curse” phenomenon by increasing awareness of the value of D2D. Assuming there is a gap in this awareness, share advantages and benefits of D2D in terms that are important and meaningful to providers and AFHTO staff.

Reflections and implications: “getting started” in QI activity

The rapid increase in voluntary participation in D2D suggests that the “get started” approach worked. The improvements in data standardization in the absence of improved functionality of EMR systems demonstrate that people were willing and able to do what they could with what they had to make a difference. The increasing amount of data contributed by each team in each iteration of D2D is also a hint that people were willing and able to build on their achievements over time. There is evidence in the literature for the value of change strategy that involves simply getting started, or what I refer to as “START-egy”. The concept of effectuation (Weick, 1985 pg. 52) suggests that doing something is a way to know something. An example of effectuation is what Janney & Dess (2004) call “immediate entry”. This involves taking small actions to provide early input to guide next steps. These next steps could include quitting to reduce the risk of becoming too tied to a lost cause that it is not possible to get out without embarrassment, a phenomenon that Bowen (1987) refers to as “escalation of

commitment”. A key element of success with these examples of “START-egy” is the focus on small changes. There is abundant evidence in the literature in support of the idea of “small” in improvement efforts. The Model for Improvement (API, 2016) identifies “small changes” as a key success factor in improvement. The model recommends multiple rapid cycles of small changes rather than single momentous changes as a way to improve outcomes in healthcare. “Small” is also a core element of movement between stages of change in Prochaska & DiClemente’s theory of change (1992). The concept of ongoing cycles of small changes is consistent with the view of change as continuous holds that organizations are “always, already changing” (Tsoukas & Chia, 2002) in small ways as they make sense of the pressures they are facing. In other words, ongoing small changes are the normal state of organizations, not just a potential strategy for achieving externally imposed change.

It is clear that D2D was able to overcome inertia related to measurement and thus acted as a disruptive artifact to stimulate change. This success might be setting up the next challenges for D2D. The first challenge is supporting sustained participation. About 85% of members have participated in at least one iteration, but no more than about 60% have participated in any one iteration. This suggests that members were able to “get started”, but did not “keep going”. Perhaps the supports that helped them get started were not helpful in supporting ongoing measurement. Evidence regarding the need to support sustainability (that is, keeping going vs getting started) includes the concept of celebration. There is very little evidence of any kind of celebration in AFHTO’s experience of D2D. Perhaps this is not surprising in an environment which prizes overachievement, competition and winning. In such a context, simply getting started might not be considered a success and therefore not worth celebrating. Nevertheless, the paucity of celebration is somewhat surprising given the emphasis on relationships and the known

role of celebration in building relationships. Celebration may be particularly important to sustaining momentum. Celebration is one of the core elements of the “control” phase of the 6Sigma approach to improvement (Tayntor, 2007). It is a mechanism of reinforcement of desired behaviours. Celebration is closely tied to the concept of “fun at work” (Plester in Owler et al., 2010). Theories of planned change such as those described by Kotter (1995), grey literature and human resource manuals abound with recommendations about the value of celebration, praise and reward (Owler et al., (2010) provide a few examples).

The absence of celebration may also be tied to the lack of progress in actual performance. The change in measurement activity was not accompanied by changes in performance or even beliefs and attitudes about quality. This might not be such a surprise given the relatively short 3year tenure of the initiative to date. The member-based steering committee felt that 5 years was a more reasonable timeline than the 3 years that D2D had been in place to see changes in performance. This decision reinforced the idea that it may be too early yet (at time of writing) to draw conclusions about the impact of getting started with D2D on improvement activity. Time will tell. In the meantime, it is worth considering other possible contributing factors to the lack of movement in performance. One might be the single-minded focus of D2D on measurement. The Model for Improvement (API, 2016) includes measuring as a core activity. However, other activities must follow measurement to result in improved outcomes. These include identifying, implementing and testing the impact of small changes on performance and then providing feedback to those involved in the process to allow them to further reflect on and adjust their approach as necessary to achieve even better performance. Another possible factor in lack of progress in performance is the “get started” philosophy of D2D. D2D was characterized by a welcoming, non-judgemental “come as you are, do what you can” spirit. In contrast, the concept

of improvement is inherently judgemental in that it suggests that the current state is not good enough. The phrase “quality improvement” itself can therefore be off-putting because of the perceived judgement it carries. Ironically, AFHTO physicians (like others) are very overt about identifying competition with their peers (a highly judgemental activity) as one of the biggest drivers for improvement in physician performance. The challenge for D2D is that the absence of judgement was intentional and instrumental in getting started. At the same time, the absence of judgement might be making it harder to maintain momentum after having overcome the initial inertia related to measurement.

All of this suggests that “keeping going” with D2D means different supports are needed for different activities than those involved in “getting started”. Suggestions include the following:

- Get started at keeping going: Find innovative ways to incorporate the successful “get started” approach of small, easy changes into strategies to support continued participation (that is, keeping going). Possible options include reframing the next steps as something else (for example, improvement) and applying the same “get started” techniques to what is effectively positioned as a “new” activity, rather a continuation of a previous one. Keep all potential changes small, both in perception and actuality.
- Celebrate: Celebrate small successes on an ongoing basis in ways that are considered fun and meaningful to those involved. Build on evidence of “organic” fun (Owler et al., 2010; Tews et al., 2014) to celebrate signs of progress with measurement and improvement on an ongoing basis. Focus particularly on teams and individuals in the action and maintenance stages of change. In these stages, the desired behaviour is

actually happening. Support in the form of reinforcement (e.g. reward, celebration) is required in these stages to prevent relapse away from the desired participation in measurement and improvement.

Summary

Reflections on the experience with D2D in the context of published literature suggest actions for next steps with D2D. The importance of relationships was observed in the experience with D2D and affirmed in the published literature. This suggests that next steps continue to focus on relationships and build on the success to date in this regard. As anticipated in the research framework, attention to conversations, particularly with physicians, is warranted to nurture and further extend the strength of relationships already established through the experience with D2D.

Observed patterns of help-seeking and self-reliance were echoed in the literature on behaviour change. When seen through the lens of theories of behaviour change, the experience with D2D suggests that further interventions be designed more deliberately to match the stage people are in with respect to change. The paradoxical help-seeking behaviour also suggests increased attention to the role of ambivalence in the experience with D2D.

The observed diversity in perceived priority of D2D may be partly explained as a gap between espoused theories and theories-in-use among AFHTO staff and members. Lack of truly informed participation in D2D may also be a factor. This suggests that future work with D2D include efforts to build awareness of what beliefs regarding measurement are truly held in common among members. This includes communicating with particular attention to understanding the extent to which the formally articulated values of the organization are actually driving operational decisions among members.

Finally, the success of the “Get started” approach to measurement was consistent with organizational change literature. It was also consistent with the research framework which posited that conversations are an intervention (Macpherson et al., 2006). The intentional focus on *getting started* with *measuring* had gratifyingly positive impacts on these specific behaviours areas, but little impact on *keeping going* and actually *improving performance*. There may be value in extending the approach of small, rapid-cycle changes that was successful in getting started to the challenge of keeping going. Celebration of progress might be a particularly useful strategy to consider in next steps with D2D.

Table 6-1 summarizes the implications for action emerging from this critical reflection on the experience to date with D2D. It represents the starting point for Chapter 7, which presents the evidence from operational documents regarding the implementation of these actions and the evidence from interactions with members and stakeholders regarding the impact of the actions taken.

Table 6-1: Themes and actions emerging from the review of data in Phase 3

Theme	Actions
Relationships	Nurture relationships Maintain attention on conversations Continue inviting providers to participate Build cadre of physician champions
Help-seeking and self-reliance	Consider stages of change Explore ambivalence
Diversity of perceived priority of D2D	Explore theories in use regarding measurement and improvement Communicate the value of D2D more effectively
Getting started at QI	Get started at keeping going Stay small Celebrate

Chapter 7 Phase 4: Response to 4th iteration of D2D (final learning phase)

Preamble

This chapter describes the third phase of this action research project, which is the experience with measurement and improvement following reflection on the initial experience with D2D, the artifact in this study. It focuses on the extent to which the actions suggested by the initial experience were implemented and whether they made a difference in the experience with measurement and improvement. Like the description of the initial experience, this chapter is essentially the answer to the first question in the critical reflection framework described by Rolfe et al. (2001): “What happened?”. As discussed earlier, my role as lead for the program was often to initiate but not complete the action. There were few actions that I took completely independent of others. Nevertheless, for the sake of tracking the fate of the suggestions emerging from the earlier iterations, I take ownership of many of the actions in this chapter even though, in actuality, attribution of action to specific actors was much less clear and of even lesser interest. This chapter also includes brief summaries of attempts to make meaning of the individual observations, which are the focus of reflection in Chapter 8. The areas of observation described in this chapter are: changes made to D2D based on initial experience, performance on D2D indicators, and conversations among AFHTO staff, members and external stakeholders. The template for the analysis of data for this chapter was based on the actions that emerged from reflection on the initial experience in the context of literature. The data sources for this chapter include operational documents (e.g. minutes oversight committees) as well as performance on the D2D report and email conversations, as in the analysis of the initial experience with D2D. This chapter, therefore, makes two contributions to the action research project: it outlines which

actions I took in response to the earlier iterations of D2D and the impact (if any) on the experience with measurement and improvement.

Changes to D2D based on initial experience

There were several changes suggested for the 4th iteration of D2D based on the experience with the first 3 iterations (see

Table 6-1). When analyzing the data, it became apparent that I implemented some of them either as proposed or in a modified form. There were some recommendations for action that I did not follow and some actions I took that had not surfaced in the reflections on the earlier iterations. These changes are summarized below.

Relationships

As suggested by experience with the initial iterations of D2D, there was continued attention to relationships. I made a deliberate attempt to more explicitly reference member input in any decisions regarding D2D. I encouraged my team to include phrases like “*designed according to what QIDS Specialists have been generating*” or “*in response to feedback from members*” in communication within and beyond AFHTO, where this was the case. I also continued to be directed by member input in operational decisions, abandoning a plan to use an existing mentoring network as an educational tool when members said they might “*feel vulnerable about having a peer teach them ... in front of their staff*”. Also, I retracted advice to teams to leverage the data extraction and summarization processes in D2D for mandatory reporting. I did this because members said this might blur the lines between voluntary and mandatory reporting and thus end up “*breaching the precious trust that teams have in AFHTO*”. Related to that, I also changed the positioning of tools to improve access to data as a resource for program planning (which was of interest to members) rather than a tool to make mandatory reporting easier.

As recommended, conversations continued to be a focus, with ongoing measurement of conversation frequency in D2D surveys. In the fourth iteration of D2D, I chose to highlight the increasing frequency of conversations in membership communications and external (that is, public) presentations.

I also acted on the suggestion to expand “invitation” as a strategy to encourage involvement. I and my team conducted explicit, personal outreach to teams without QIDS Specialists support and teams from whom there had not yet been signs of interest or progress in contributing to D2D 4.0. In addition, we invited clinicians (IHPs in particular) to participate in a variety of activities to inform and support improvement. These included inter-professional knowledge translation and exchange sessions on motivational interviewing as a clinical improvement strategy and focus groups with inter-professional healthcare professionals to define their role in improvement.

I enlisted AFHTO staff and selected committee members to further build the cadre of physician champions. We began direct personal outreach to medical leads of teams with messages co-signed by physicians in Board leadership positions, to encourage their involvement in contributing to and using D2D. We contacted physicians with whom we had relationships to get clinical input regarding definitions of D2D indicators. I also invited physicians in leadership roles in AFHTO (members of committees, Boards) to participate in a study aimed at better understanding physician workload.

Help-seeking and self-reliance

The pattern of behaviour regarding problem-solving suggested that interventions to advance D2D be matched with the stage of change people are in. Without direct reference to this suggestion, I redoubled efforts to reduce workload associated with D2D as this was repeatedly highlighted as a barrier. For example, I pushed the timeline for D2D 4.0 out by several months to give members more time to prepare. This was appreciated. However, my effort to reduce workload by more closely aligning D2D with mandatory reporting processes was rejected. Again, without overt reference to behaviour change theory, I highlighted the extent to which the

“pros” of participating in D2D (e.g. better quality data, improved strength of AFHTO’s advocacy for members) might outweigh the “cons” of the extra work. Although some of these efforts were appropriate for specific stages of change, I did not consciously consider that, nor target them, to the relevant people. Neither did I explicitly explore ambivalence in the observed problem-solving behaviour. Despite the power and control over D2D inherent my role as program lead and principal investigator, there clearly were other forces of equal or greater power affecting my decisions. My lack of action *and* lack of awareness of my inaction were surprising and troubling observations that prompted deeper reflection on my role and the impact of power on the experience of change associated with D2D (see Chapter 9).

Diversity of perceived priority of D2D

The suggestion coming out of the reflections regarding the diversity of perceived priority of D2D was to build awareness about beliefs and values regarding measurement and improvement among AFHTO members and staff. Other than gentle inquiry to better understand the barriers to participating in D2D, I did very little in this regard. As with my lack of action regarding ambivalence, this gave me pause for thought about the impact of power on my choices. I address this in more depth in the reflections and implications of the D2D 4.0 experience for the future (See Chapter 8).

I did, however, act in response to the suggestion to communicate the value of D2D more effectively. I and my team released considerably more communication materials through more avenues. We made a printed version of D2D available in addition to the original interactive website. I made videos about the composite measure of quality. These were released publicly along with an endorsement from a prominent physician. My team also convened an exhibit about D2D at the AFHTO annual conference and several other meetings. We designed, ordered

and distributed special “I do D2D” bags at the AFHTO conference. I and several AFHTO members made multiple presentations at premier family medicine conferences across North America. I was also accepted into research mentoring program, to which I applied to serve as an avenue for sharing the D2D story. Notwithstanding the increased amount of communication and use of different modes, the *content* of communication (that is, focus on the value of D2D) did not change substantially.

Getting started at QI

The experience with the initial iterations of D2D suggested extending the success of the “get started” approach to the challenge of supporting AFHTO members in “keeping going” and achieving improvements in performance. To that end, the governance committees for D2D developed indicators of “using D2D” as an interim step towards improved performance. These groups also focussed on variation in performance between teams, not just the mean values of D2D indicators. Another action was framing D2D as a “lightning rod” to drive more local (and therefore potentially more meaningful) efforts in data quality and measurement, independent of contribution of data to D2D. For example, my team worked with members to produce a catalogue of program-level indicators in use among AFHTO members. We described it as a tool to make it easy for teams to see what others were using for local, program-specific measurement. We were explicit that it was *description*, not *prescription*, about what indicators they should use. Comments from members indicated they wanted AFHTO to focus on the ground-up approach of D2D and the extent to which D2D was responsive to front line providers, not just the actual performance reported in D2D. In addition, the steering committee approved my proposal of a series of QI resources to help members move beyond measurement to improvement.

In terms of the recommendation to celebrate, I was able to take some small actions. The special “I do D2D” bag distributed at the annual conference was intended to recognize and celebrate the teams that had contributed to D2D. I also applied (unsuccessfully) for an award for D2D (in response to member suggestion). There was more frequent reference by AFHTO staff and in Board documents to the relationship between high quality care and lower *per capita* healthcare costs observed through D2D. I and my team kept trying to elicit “success stories” from members. I found one fascinating example of celebration of improvement at the local team level during my routine visits to the field. The team was disappointed that they were not chosen to receive a “Bright Light” award from AFHTO. They decided to create their own award for themselves (see Figure 7-1). The iconic “soft glow of electric sex” (‘A Christmas Story’, 1983) lamp they chose adds an element of fun to the spirit of celebration that inspired them. They went on to receive not one, but two, Bright Lights awards the following year.



Figure 7-1: Gift of staff of one team to leader to celebrate achievements within the team (i.e. local celebration)

In keeping with the guidance emerging from the initial iterations, D2D remained small. AFHTO members further clarified that “small” meant 12-20 indicators. While it continued to evolve, the next iteration of D2D (D2D 4.0) was deliberately and overtly very similar to D2D 3.0. This was highlighted in much of the D2D communication.

Other changes for D2D 4.0

A more ideal action research project might have ensured full implementation (or at least overt attempts at implementation) of all changes emerging from reflection on the earlier actions.

However, D2D is not and never was a research project. It was and is an operational activity of AFHTO, independent of its role as a focus for this action research. Research findings are not even remotely the only considerations in my decision-making regarding D2D. Instead, my decisions were, and are, guided by constant negotiation among the multiple stakeholders in AFHTO: members, staff, funders, external partners and patients, to name a few. The actions I (and others) take are a balance between what is ideal and what is possible and are consistent with the spirit of “do what you can” that characterizes much of D2D. Consequently, only some of the changes surfacing from the reflection on the first iterations of D2D were implemented. Further, I made several operational changes that were not based on the action research findings, but were direct responses to requests from members. These included eliminating the sign-up process for D2D 4.0 and changing the composite indicator in response to suggestions that might make it more meaningful and useful to members.

Table 7-1 summarizes the recommended changes and the actions taken (if any) on each. The remainder of this chapter examines the experience with D2D 4.0 through the lens of the recommended changes emerging from the earlier phases of the action research. The use of this lens is not intended to judge whether the “right” things were done. Instead, it is intended to facilitate understanding about the “right”-ness of the recommendations through reflection on the extent to which they were acted on and considered useful.

Table 7-1: Fate of recommended actions for D2D 4.0

Theme	Actions	Evidence of action
Relationships	Nurture relationships	Explicit references to impact of member input in D2D implementation
	Maintain attention on conversations	Increasing frequency of conversations highlighted in communication materials
	Continue inviting providers to participate	Personalized outreach to teams re: D2D and to clinicians re: improvement resources
	Build cadre of physician champions	Communication to teams via Medical Leads, endorsement by prominent physician
Help-seeking and self-reliance	Consider stages of change	No reference to change theory although some actions taken might be stage-appropriate
	Explore ambivalence	No evidence of action
Diversity of perceived priority of D2D	Explore theories in use regarding measurement and improvement	Tentative inquiry to better understand barriers for D2D, but otherwise no evidence of action
	Communicate the value of D2D more effectively	Higher volume and more modes of communication, but little change in focus on value of D2D
getting started at QI	Get started at keeping going	Focus on interim markers of progress (using data, reduced variation); introduce a series of improvement resources (as distinct from measurement resources)

	Stay small	Definition of “small” was articulated (that is, 12-20 indicators), overt limit on changes for D2D 4.0
	Celebrate	Very small signals (e.g. D2D bags)
Member-requested changes	Reduce workload/make it easier	Eliminate sign-up process and formalize bi-annual schedule for D2D reports
	Make composite measure more meaningful	Present drill-down into recognizable component measures, update component measures
	Make geographical comparisons possible	Expand D2D scope to include LHIN so teams can identify others in their regions

Participation and performance in D2D 4.0

Participation: There was no change in the number of teams participating in D2D 4.0 relative to 3.0. However, the trend of contributing more data continued. Unlike D2D 3.0, the communication materials related to the launch of D2D 4.0 made explicit reference to the breadth of participation in D2D, highlighting the very high proportion of members (85%) who had been part of at least one iteration.

Team characteristics: The frequency of conversations about measurement continued to increase. In addition, more teams identified physician champions. Other measures of team characteristics described in earlier iterations were not available for 4.0 because of the low response to the survey.

Performance: As with previous iterations of D2D, there were several areas of focus for performance measurement. EMR data quality decreased in D2D 4.0, as did the number of teams

contributing data for this measure. However, for teams contributing to the measure in both 3.0 and 4.0, performance increased. Average scores on individual and composite quality indicators and cost were unchanged from D2D 3.0. However, there was a slight decrease in variation in the scores between members.

Conversations

The Level 1 codes of the coding template for the conversations regarding the 4th iteration of D2D were the proposed changes emerging from the experience with the first 3 iterations. The final template of Level 1 and Level 2 codes emerging from the data with representative text is shown in Appendix 3. These data were aggregated into themes, loosely organized according to the areas of action recommended by review of earlier iterations of D2D. These observations are also considered in the context of the quantitative data presented above to support the identification of themes from the combined qualitative and quantitative data describing the experience of D2D 4.0.

Experience with actions aimed at addressing Relationships

The data illustrate a positive impact from the continued attention to relationships. One new theme in the experience of D2D 4.0 was the protective tendency around relationships with members arguing against proposals they felt carried “*risk in breaching the precious trust that teams have in AFHTO*”.

The conversations among members also illustrated local improvements in physician engagement. This story from one member conveys the changing culture, the perceived importance of conversations and excitement about all that:

“The most exciting thing is the momentum ...we now have two physicians that have permitted our students to code their records. One of these physicians last year had blamed us for

not being able to get patient lists for him and refused to talk us LOL. What a change in a year. Yesterday another physician that I'd describe to-date as ‘an island’ actually agreed to pilot an [EMR data capture tool] in exchange for us updating some forms that really bug her. I bet conversation time with her yesterday was 20 minutes!”

However, this was balanced by observations about teams who “*had to hunt down the necessary clinicians*” and reported a continued “*reluctance of physicians to provide these data*”. This suggests that physician engagement remains a challenge.

Personalized invitations seemed to be well received, if the actions taken in response are any indication. For example, teams specifically invited by their local peers agreed to contribute data to D2D for the first time. Another suggestion that personalized invitations were appreciated was the disappointment of some members who “*were wondering how come they had not heard of the upcoming AFHTO event*” and therefore felt excluded.

The data emerging from conversations is consistent with quantitative data of persistently high participation that suggest that the strength of relationships is increasing. This may be due to the continued attention to them through overt acknowledgement of input and the explicit reference to the increasing frequency of conversation in quantitative summaries of D2D data. Protecting the relationships and personally inviting more and more people to participate in the process may also be contributing to the increased strengths of relationships. This represents a continuing reason for celebration for AFHTO as well as a cautionary note about the risks of investing heavily in relationships. These risks can include decreased openness to discordance and disagreement and a risk of members feeling left out if limited AFHTO staff resources interfere with reaching all members using a personalized approach.

Experience with actions aimed at addressing Help-seeking and self-reliance

Qualitative data regarding help-seeking and self-reliance confirm impressions from operational data that little action was taken to match interventions to readiness to change. As with earlier iterations of D2D, there is little in the quantitative data regarding help-seeking and self-reliance, other than the persistently low response rates for surveys, which seems consistent with low engagement in self-help behaviour. The qualitative and quantitative data suggest a need to do more to include myself as the researcher among the researched. My inaction on my own recommendation is worthy of reflection, especially since the recommendation itself was intended to make it easier for others to take their own actions. It suggests I am mirroring the tendency of members to *not* take actions that are known to be potentially effective. This apparent fractal nature of the problem being investigated is addressed in more detail in Phase 5 (Chapter 8).

Experience with actions aimed at addressing Perceived priority of D2D

As with earlier iterations of D2D, the only quantitative data relevant to the perception of priority was the participation rate, which persisted from the previous iteration. The remainder of this section about perceived priority is therefore based on qualitative data alone. Members noticed the increased volume of communication material, commenting that there was “*Plenty of material here for proselytizing!*” and also that the “*videos were well-received*”. They were pleased with the introduction of a print function for D2D and subsequently observed “*more clinician engagement with the printed report*”. Nonetheless, communication gaps remained. Members reported continuing “*low awareness of D2D among clinicians*” and “*a particular [unmet] need to share the encouraging news to support teams in moving from measurement to*

improvement”. Unlike earlier iterations, members made numerous suggestions to improve both the process and content of communications, as illustrated by the comment below:

“A greater emphasis must be placed on the importance of D2D influencing government and how the data is of value to team. Members must be able to see the value in participating is greater than the burden of getting the data from the EMR with emphasis on telling members the value of correctly entering data in their EMR would greatly decrease the burden of extracting the data for D2D”

Other suggestions included increased profile of “*good news*” and framing opportunities for improvement in a positive light: “*you are good – and we can do better!*”, as well are more attention to the extent to which confidence in using D2D was increasing.

Despite the apparently heightened efforts related to communication on the part of AFHTO as well as members, there was no evidence of attention to the underlying differences in perceived priority of D2D. There was more exploration of barriers to D2D (mostly related to workload) and discussion of possible solutions. Beyond that, there was no overt recognition of the persistent differences in perceptions regarding the priority of D2D. The pattern of adjusting and extending the effort directed to the same communication activities (albeit through some new vehicles) with virtually no attention to the underlying differences in priorities is an example of single-loop learning. Further, the failure to explore the divergence in perceived priorities that was necessitating the communication efforts is a missed opportunity for double-loop learning. Since double-loop learning is uncommon in organizational behaviour (Argyris, 1999), AFHTO is not unlike many other organizations. That was not surprising. However, I was disappointed to notice that, even with my heightened interest in action learning, I also failed to take up the challenge of double-loop learning in this situation. This bears further examination, considering

its potential impact on the success of D2D. My reflections on this are explored in more detail in Phase 5 (Chapter 8).

Experience with actions aimed at addressing Getting started

As in earlier iterations of D2D, there was awareness and appreciation of the “get started” approach. Some felt that the “*Biggest priority for AFHTO is to promote the approach*” of D2D. Members noted that there is great value in knowing that “*D2D is NOT an all or nothing exercise*”. They accept that “*It’s not perfect, but getting better*”. Members accepted that it was reasonable to not have seen improvement in performance at this stage in the initiative. The fact that “*more teams are not just participating in D2D but actually reviewing the report*” is an example of the evidence cited to defend the perception that AFHTO is making progress with D2D even without changes in performance. There was also quantitative data suggesting increased attention to EMR data quality, which is consistent with the sense of members that they were making incremental progress towards improvement.

There was continued interest in exploiting the ability to compare to peers “*to help other teams decide who they want to talk to help them improve*”, effectively reducing variation between peers. Reducing variation was explicitly identified as an important goal for D2D even as it was acknowledged that “*it is different work from measurement*”.

There was increased profile of the importance of IHPs, with a sense that “*IHPs and QIDS Specialists are well positioned in ideas for moving forward for initiatives on improvement*”. IHPs were considered to “*have a vested interest ...because they are usually asked to contribute data [that] reflect progress on the programs they lead*”.

Finally, there was some suggestion of movement regarding celebration. For example, there were small hints of celebration in the form of compliments: “*This is SUCH great stuff!?*”

and “*Your contribution is undoubtedly remarkable!*”. In addition, members identified achievements of members (vs AFHTO staff) that were worthy of more celebration, suggesting that AFHTO “*Deliver good news messaging about D2D to internal and external stakeholders.*” *There are some real success stories coming out of the work the QIDS program is doing*”. Another suggestion included dropping the adjective ‘success’ from the invitation to share stories to overcome “*the hesitation that it would seem presumptuous to claim what a [team] is doing qualifies as a success*”. These kinds of comments from members were new for D2D 4.0.

The experience of members with D2D 4.0 suggests that the “get started” approach continues to be appreciated and effective. Supporting it further involves more attention to demonstrating progress with “getting started”, as distinct to “being finished”. This could also better support identification of reasons for celebration. The identification of reducing variation as an interim goal for improvement, especially with the active participation of IHPs has implications for how the organization proceeds with its plans to offer resources aimed at improvement, especially given the focus on quality improvement staff to date.

Summary of Phase 5

The observations based on quantitative and qualitative data were considered together to identify themes in the experience of D2D 4.0. Because the goal of this phase of the action research was to examine the response to actions recommended out of consideration of earlier iterations, the themes were organized according to the recommended actions. Regarding the actions about relationships, the data showed increased strength of relationships as well as positive impact of personalized invitations and physician outreach to peers. The recommended actions related to help-seeking and self-reliance were, for the most part, not implemented, raising the potential need for more reflexivity in this action research project. Similarly, my lack of

engagement (and that of my organization) in “second loop learning” about perceived priorities suggests reflection on factors contributing to efforts to expose my assumptions as a researcher as well as those of my organization. On the other hand, the increased awareness of members of the need for communication (as evidenced by their suggestions to improve it) suggests progress with respect to understanding the importance of D2D. Finally, the experience of D2D 4.0 illustrated continued effectiveness of and appreciation for the “get started” approach that had embodied D2D from the outset. Member suggestions to leverage this approach could guide action for subsequent iterations of D2D. These observations are summarized in Table 7-2 for deeper reflection in Phase 5 (chapter 8).

Table 7-2: Dimensions emerging from review of data in Phase 4

Level 1	Aggregate dimension/theme	Level 2
Getting started	Celebration	actual celebrations; compliment; stuff to celebrate; approach; stories vs success stories
	Increased use	make it easier; alignment with requirements; good enough now; measure measurement; tools; stay small; demonstration of increased use; using data; facilitate use
	Reduce variation	feedback mechanism; reduce variation; set expectations
	Use IHPs	program planning indicators; use IHPs
Help seeking/self reliance	Fractal nature of problem being investigated	stage-specific interventions; contemp; precontemp; prep

Perceived priority of D2D	Absence of 2 nd loop learning	examples of success in communication; not there yet
	Incremental progress with communication	develop more content; messaging suggestions; process suggestions
relationships	Increased strength of relationships	incorporating member input; example of input; impact of input; seek input; nurture relationships; building relationships; leverage relationships; protecting relationships
	Invitations	inviting participation
	Physicians	improvement in engagement; little physician engagement; strategies to improve engagement; build phys champs

Chapter 8 Phase 5: Reflections on the 4th iteration of D2D (final learning phase)

Preamble

Phase 5 involves a deeper reflection on the experience of AFHTO with the 4th iteration of D2D as described in Chap 7. It addresses the question “So what?” in the critical reflection framework of Rolfe et al. (2001). The experience with the 4th iteration was described through the lens of the actions emerging from the earlier iterations. Specifically, these included actions to address relationships, help-seeking and self-reliant behaviour, diversity of perceived priority of D2D and the value of getting started in measurement and improvement (see

Table 6-1). In addition to describing what happened with each of the actions, Chapter 7 also introduced reflection on the observed experiences. As with Phase 3 of the action research project, Phase 5 considers these reflections on individual observations in a more collective way and in the context of evidence from the literature. These reflections refine the themes emerging from the experience with the 4th iteration to facilitate sense-making of the overall action research project. Since the D2D initiative continues beyond the scope of this action research project, these reflections were considered in the design and implementation of the fifth and subsequent iterations of D2D. However, this chapter does not explore recommendations for action in detail because the action research project concludes at the end of the 4th iteration of D2D.

Reflections

Relationships

The increased strength and appreciation of the benefits of relationships among members was one of the goals of suggested action coming out of the earlier iterations of D2D. This may simply reflect the natural development of the social capital life-cycle. The observed pattern in relationships is consistent with social capital theory, which describes an initial focus on “bonding” to build internal connections followed by an increased interest in “bridging” to extend relationships to others for concrete mutual benefit (Roberts & Coghlan, 2011). My efforts to reference member input as a driver of decisions appears to have been appreciated. This is consistent with literature that suggests sustained meaningful engagement depends on clearly demonstrating the impact of input from participants (CFHI, 2013), something that is not only respectful, but also builds trust and ongoing interest in collaboration. Dorazio (2014) observed that demonstrating that input has been heard is a key driver in citizen engagement. Being overt about the impact of contributions was identified as a core element in several engagement toolkits

(New Zealand Internal Affairs, 2015; CHFI, 2013; Government of Canada, 2016). It therefore seems reasonable to conclude that overtly referencing the extent to which member input was driving decisions in D2D 4.0 contributed to the observed increase in strength of relationships. This is positive. However, Janis, (1973) suggests that groups with an intentional focus on collegiality and positive relationships must be vigilant to avoid loss of critical input and developing of Groupthink. This could perversely decrease the ability of the organization to hear input and thus interfere with the organization's ability to learn from its experience.

The role of physicians in influencing their peers is well-established in the literature. For example, in a review of dissemination strategies to change practice, Kanouse et al., (1995) described peer influence as highly effective way to influence physician behaviour. More recently, Ivers et al. (2014) listed using a “trusted source” to communicate feedback to facilitate physician practice change. More locally, the role of physicians in influencing their peers was a core part of the decision to fund the recruitment of more than 70 physician leaders to facilitate the implementation of the primary care transformation agenda in each of the newly-defined sub-regions of Ontario. The pattern of physician influence on physicians observed in the D2D initiative is thus consistent with knowledge about physician learning.

The effectiveness of specific invitation (vs broadcast email) was gratifying. It is consistent with early advice from QIDS Specialists who suggested (and demonstrated) that they were more willing to contribute to “round table” discussions if they were explicitly asked by name to do so than if the invitation was non-specific. This is consistent with engagement literature across many sectors. The evangelical Christian movement has long identified personal invitation as one of their most important recruitment strategies (Stetzer, 2014). Personalization of communication is widely identified as a valuable enabler for success in the field of marketing

(BakerGoodChild, n.d.) and engagement in membership organizations (Jenkins, 2016). Sezginalp (n.d.) suggested that direct contact could build the commitment of members to their association. Possibly because of the strong consensus that personal contact works to increase participation, there does not seem to be much exploration about why that is so. It may be worth addressing this gap. For example, QIDS Specialists reported that they responded more readily when asked directly to contribute because being asked (instead of volunteering) made them feel less presumptuous about sharing their experience. This was contrary to my belief (and the overall nature of D2D) that the best approach to ensuring participation in anything was “voluntary”. It was also counter to the assumption that direct invitation to share might be perceived as an imposition. These surprising reasons for QIDS Specialists response to invitation could possibly have predicted the disappointment some members felt about being left out of personal invitations to participate in other activities. This reaction needs to be considered carefully as it represents a potential and previously unknown risk in extending personalized invitations. It also highlights observations by Tschirhart & Gazley (2014) that the literature on how membership associations work remains sparse. The mechanism for the effectiveness of personal direct invitations to participate is therefore worth further exploration even while it remains an effective strategy for D2D.

Implications:

In the ongoing attention to building and strengthening relationships, effort should be made to reduce the risk of “group think” and the loss of appetite for critical thinking and input from front-line providers.

It is also important to explore the mechanics and potential risks of personalization in building engagement. Clearly, direct contact is effective in increasing participation. However,

when scope of personalized invitation is limited by scarce resources (as is the case with AFHTO), there may be a risk that those not reached might feel left out and perversely *less* engaged by a strategy intended to increase participation.

Help seeking and self-reliance

The recommendation to design and implement stage-specific interventions to support change according to the readiness of individuals for change was well-grounded in literature. It was consistent with multiple models of self-change such as those described by Prochaska & DiClemente (1992), Schwarzer (2008) and more historically, Rosenstock (1974). It also resonated with my experience in using this approach to support change in other settings. However, I didn't act on it. This suggests that I need to expand the focus of the problem being investigated here to include my failure to act on viable strategies. How uncomfortable! This fractal-like phenomenon is both justification of, and fodder for, reflexivity in research. It seems to be a clearer example of foundational premise of this doctoral program: "the problem is in me and I am in the problem" than might be seen in other endeavours. It is commonly understood that it is human nature to see more clearly the failings of others than our own. It is equally well accepted that the only behaviour any one person can change is their own. This irony points to the value of reflexivity: if I can see more clearly what the problem is by looking at others and then recognize it in myself, my reflections about what I can or will do about it might help inform actions I can take to support others. Having said all that, reflexivity is still uncomfortable as it removes the distance between the researcher and the researched and traps me in the "swampy lowlands" (Coghlan & Brannick, 2014, pg. 4) where I, like those I am with, am unable to examine my own assumptions.

There were probably several factors involved in my response to recommendation for stage-specific strategies. First, even though I made the recommendations, I don't think I was convinced that I needed to do anything differently. Because I consider myself to be skilled in applying change theory to my work, I believed that I intuitively was doing so and therefore didn't need to make any specific, overt changes in my approach. The evidence suggests that this is partly true in that I did undertake some strategies that were more appropriate to particular stages, and I was able to recognize evidence suggesting the stages people were in. However, the evidence is also clear that the problem of untargeted interventions persisted, suggesting that overt change in my approach was clearly necessary.

On reflection, another factor might have been a fear of losing face with members. I am aware that behaviour change theory (particularly the Stage-of-change model) is generally not widely understood. Possibly because of that, behaviour change strategies are also considered to be manipulative, a term which accurately describes the skilful application of theory to support change, but which is often perceived to be negative (Mitchell, 2015; Little & Girvin, 2002). I may have unconsciously chosen not to be overt about this approach for fear of being judged as manipulative by the members I was trying to serve.

Similar reflections apply to my lack of action regarding ambivalence. The driver here might have been more related to the fear of exposing assumptions and, more specifically, the combative defensive reactions of others if, and where, their assumptions might have been exposed (Argyris, 1999). Regardless, the awareness that there are drivers working against a decision that I ostensibly believe in (having made it myself based on my observations and my understanding of the literature) underscore the importance of exploring ambivalence not only in my own thinking, but also among AFHTO members.

Implications:

The value of these reflections is not in the opportunity for self-flagellation, but rather to build my capacity for empathy. Empathy is a driver of effectiveness in qualitative research (Lugosi, 2006). It is also an important attribute of effective leaders (Randall, 2004) and key to changing our world for the better (Pava, 2008). The value of these reflections is also their ability to provide practical insights about other ways to support others in changing their own behaviour by starting with my own. For example, more awareness of the need for change (that is, I'm not as skilled as I think) may help support me in changing as well as encourage others to do the same. To further illustrate the fractal nature of reflexivity, the solution emerging here is, in fact, a stage-specific intervention that is appropriate for people in pre-contemplation, a group that is fairly prevalent among AFHTO members. Further, awareness of my own ambivalence creates awareness of the possibility that the same phenomenon is present among the members of the organization I serve. This could help build my empathy and potential effectiveness in supporting them in changing their behaviour.

Perceived priority of D2D

The increased volume and breadth of modes of communication could be seen as a positive change over previous iterations. In a summary of best practice in feedback and audit, Ivers et al., (2014) recommend multi-modal communication is always a good idea. Like many organizations, AFHTO has received advice from experts to extend the range of modes used for communication. The change in communication patterns, coupled with the increased involvement of members in attempting to direct communication, could therefore be considered positive progress and therefore reason to celebrate.

However, the focus on quantity rather than quality of communication bears more examination. My inaction regarding gaps between espoused theories and theories in use about measurement and improvement is another opportunity for reflexivity. As noted above, my own behaviour echoed the problem I was investigating in this action research: I did not follow some of my own advice regarding next steps with my own work. I had my reasons. Based on knowledge accumulated in my role as an “intrapreneur” (Björkman & Sundgren, 2005), I feared that the very suggestion that D2D was not treated as a high priority might not be well received. Whether that was a valid fear or not is not the point. According to Argyris (1999), the ability to explore ones underlying assumptions is a key component of “double-loop learning”. Therefore, the important learning from my awareness of the gap between my action and my expressed beliefs is not so much the reasons for the gap, but the fact that there is a gap. Exploring the drivers for the gap might help me become more curious, empathetic and effective in supporting similar hard work among AFHTO members and stakeholders.

Implications

The apparent progress in communication may be worth rewarding with more celebration to further encourage these changes in awareness and hopefully, eventually, behaviour. Addressing the differences in perceived priorities of D2D may be disruptive in that it would involve double-loop learning and the discomfort associated with it. It may be worth re-examining the need and potential value of better understanding the beliefs about D2D relative to the discomfort that the search for this understanding might generate.

Getting started

Increased use of the D2D report is an example of a quality improvement activity. Framing it as a step towards improved performance is therefore consistent with the research

framework that suggested there were multiple steps between reporting performance and seeing improvement. It is also consistent with quality improvement literature. The Canadian Foundation for Healthcare Innovation (with a mandate to support innovation in better, more efficient patient care) implemented a strategy to address the question: “how well do healthcare providers … actually use data?” (CFHI, 2014). America’s Agency for Healthcare Research and Quality created a “playbook” to support improvement in healthcare that outlines the different processes and enablers involved in using data vs collecting it (that is, creating reports) (Korsen et al., n.d.). Beyond healthcare, Derrick-Mills (2015) reports on an extensive exploration of factors contributing to the use of data to spread this behaviour further in the education system.

The role of variation is prominent in quality improvement theory (API, 2016; HQO 2013). The importance of statistical process control tools in improving quality is based on the premise that variation beyond chance is a signal for the need to improve (API, 2016; HQO 2013). Health Quality Ontario (2013), identifies reducing variation as one of the key strategies in improving quality. The goal of comparing performance to peers is to achieve the same levels of quality as others are doing. The increased ability to help teams see “where they stack up” relative to their peers was one of the most valued perceived benefits of D2D. The perceived feasibility of achieving of what others are already achieving can be a strong motivator to improve at least to the level of peers. In this way, peer comparison can serve as a strategy to reduce variation. Reducing variation by enabling peer comparison also leverages the competitive nature of teams, and physicians in particular. Finally, setting a goal of reducing variation (rather than achieving a specific target) is consistent with the wide-spread sense of

“good enough” that characterizes the experience of AFHTO with D2D. It implies that the level of performance already being achieved by most teams (that is, average performance) is reasonable.

Accepting the current level of performance might seem paradoxical to a vision to improve performance. IHI (n.d.) cautions against being satisfied with meeting minimum standards. Berwick (in Hospital Research and Education Trust, 1992) bemoans the effort extended by healthcare providers to prove they are good enough, rather than directing those efforts to have excellent performance. However, there is emerging consensus in the field of quality improvement about the need to fail early and fail often (Martin, 2013). This is not a new idea. In the late 1980’s, a highly successful software development strategy called “worse is better” (Gabriel, 1992) emerged. It embodied an old English aphorism that perfect is the enemy of good by intentionally focussing on products that might not have as much functionality (in other words, “worse”), but were more practical and usable (supposedly better). Extending this philosophy to performance by accepting worse performance (that is, lower than ideal or target) could constitute an easier and less intimidating starting place that could better enable motivation and continuing effort to improve. Introducing the artifact of D2D was more gentle than creating a crisis but clearly it was still sufficiently disruptive to generate conversations and thus support change.

The interest in focussing on IHPs to enable improvement is emblematic of the art of the possible. There is widespread perception that physicians are not interested in QI activity and that there is nothing anyone can do about it. Assuming the perception is true, working with IHPs is more likely to be successful than trying to get started with physicians. On the other hand, if the assumptions about physician interest are not true, it still is worth starting with IHPs as this might

trigger more obvious signs of physician engagement. This could not only increase participation in QI, but also contribute to learning about how to assess and achieve physician engagement. In any event, starting with IHPs is rational because the potential value of IHPs in improving performance is neither unknown or unpredictable. The Conference Board evaluation of Family Health Teams (2014), a very credible report in my organization, showed the clear advantage of IHPs in achieving high quality of patient experience as well as clinical outcomes of primary care. Focussing QI supports on IHPs is therefore an example of acting to learn. The necessity of having to defend the focus on IHPs is ironic, given that IHPs are the key element that distinguish primary care teams from other models of primary care. The fact that working through IHPs was not the obvious default strategy of an organization aiming to demonstrate the value of teams is another example of the gap between espoused theory regarding the value of teams and theories in use in the organization. It is also an example of the power of the artifact of D2D in generating conversations that disrupt by surfacing topics that were previously undiscussable or at least undiscussed.

The value of celebration in achieving desired and sustained changes in behaviour has already been discussed in Chap 6. The decision to celebrate (or not) can be viewed as a behaviour and thus can be considered through the lens of behaviour change theory. For example, in the Prochaska & DiClemente model of change (1992), the first stage of change (precontemplation) is characterized by the absence of perceived need, ability or benefit of changing. The next stage towards actual change is contemplation, in which there is a sense of the need to change. In this case, the increase in awareness that there is something to celebrate may be considered to be analogous to movement from precontemplation (where there is no sense of anything to celebrate so therefore no need to do so) to contemplation. Therefore, while there

was no evidence of a change in the extent of celebration or “fun”, the increased awareness of achievements worth celebrating could be viewed as progress. This is another example of the gentle disruption associated with D2D in that it raised awareness and generated conversations that had not previously been part of the organization’s experience.

Implications

Conclusions about the success of D2D need to consider the evidence of movement on interim markers of progress in addition to tracking the desired outcome of improved performance. A “get started” approach, by definition, is not intended to achieve the final outcome. Therefore, markers of success in getting started (as distinct from markers of success in reaching the finish line) need to be articulated. While process or markers of interim progress are important in any change initiative, failure to do so in an approach that is deliberately focussed on getting started may be particularly crippling. The absence of such markers could hamper success by overlooking opportunities to reward (celebrate) progress towards desired behaviour and outcomes. Drawing conclusions about the effectiveness of the approach to date without considering the incremental progress may also distract attention towards developing new approaches, rather than leveraging the momentum, subtle as it is, of the interventions already underway.

IHPs are the distinguishing element of team-based primary care relative to other models. Efforts to demonstrate the value of this model may be well advised to leverage this unique human resource. D2D did not start with IHPs. It started with QIDS Specialists because it was an activity of the QIDS program and therefore could most easily start there. It does not need to stay there. Attention to the role of IHPs is advised to help AFHTO keep going with measurement, now that they have had success in getting started.

Novel measure of quality

In this closing set of reflections on the D2D journey, I feel it is important to point out something that is missing from all of the Results chapters. Nowhere in any of the qualitative data is there any evidence of the importance of the novel composite measure of quality in attracting interest of members in D2D. This contrasts with the initial premise of the D2D intervention which assumed that this approach to measuring quality would make a difference in participation. There is no question that the composite measure of quality remains important to AFHTO, especially in light of its usefulness in demonstrating the relationships between higher quality primary care and lower healthcare system cost, as predicted by Starfield. However, it is equal parts fascinating and humbling to note the gap between my expectations of its importance and how members experienced it.

Summary

Phase 5, the final phase of this action research, illustrates that the actions taken in response to the earlier iterations of D2D had at least some impact. The increasing strength of relationships is gratifyingly consistent with ongoing effort in this regard as well as with various literature regarding the importance of acknowledging input, peer-based physician education and direct marketing. However, it also raises the need to be attentive to the risk of developing groupthink behaviour which could limit the effectiveness of D2D.

Consideration of the recommended actions regarding help-seeking and self-reliance illustrates the fractal nature of the research problem being investigated. My own failure and that of the organization to engage in double-loop learning” around ambivalence and espoused theories/theories in use is disappointing to recognize. Given that this pattern is within the norms

of organizational behaviour (Argyris, 1999), next steps might focus on exploring how to support deeper learning for myself and my organization.

The subtlety of evidence regarding incremental change highlights the importance of defining measures of interim success in a “get started” approach. More clearly recognizing progress might make it more possible to leverage the power of celebration to encourage desired behaviour. It might focus the conversations that have already started in response to the disruption of D2D onto specific action, in this instance, celebration. The interim measures of progress with getting started could therefore be seen as an extension of the gentle disruption associated with D2D. The sobering realization that the novel composite measure of quality (or any other measure, for that matter) was almost completely irrelevant to the way members experienced D2D reiterates the importance of being attentive to the process (not necessarily the content) of D2D going forward. Table 8-1 outlines the reflections on the experience with D2D 4.0 and implications for next steps with D2D. Because this is the final phase of the action research project, no detailed consideration of next steps for action is presented. However, I and/or others at AFHTO may choose to explore the implications presented here in considering the next steps for D2D since the initiative continues as an operational priority within the organization. The remaining work to be done on this action research project is a moment of reflexivity which is the focus of the next and final chapter of this thesis.

Table 8-1: Reflections and actions from earlier iterations of D2D, reflections on D2D 4.0 and implications for consideration by organization for future iterations

Experience: D2D 1.0-3.0	Actions: Based on D2D 1.0 to 3.0	Experience: D2D 4.0	Implications: D2D 5.0 and beyond
Building relationships	<ul style="list-style-type: none"> • Nurture relationships • Maintain attention on conversations • Continue inviting providers to participate • Build cadre of physician champions 	<ul style="list-style-type: none"> • Continued attention to and increased strength of relationships • Predictable effectiveness of peer influence among physicians • Effectiveness of personalization in increasing participation in D2D 	<ul style="list-style-type: none"> • Celebrate and leverage relationships but also consider risk of group-think • Explore mechanics of personalization and address risk of feeling left out
Help-seeking and self-reliance	<ul style="list-style-type: none"> • Consider stages of change in designing interventions to advance • D2D 	<ul style="list-style-type: none"> • Lack of action suggests fractal nature of the research (that is, researcher has same challenges as the researched) 	<ul style="list-style-type: none"> • More overtly include researcher among the researched

	<ul style="list-style-type: none"> • Explore ambivalence regarding problem-solving behaviour 		
Diversity of perceived D2D priority	<ul style="list-style-type: none"> • Explore theories in use regarding measurement and improvement • Communicate the value of D2D more effectively 	<ul style="list-style-type: none"> • Increased activity in absence of change in focus emblematic of absence of second-loop learning • Increased member awareness of need to improve communication 	<ul style="list-style-type: none"> • Explore and build organizational support for second-loop learning
Importance of getting started at QI activities	<ul style="list-style-type: none"> • Get started at keeping going • Stay small • Celebrate 	<ul style="list-style-type: none"> • Evidence of using D2D identified as progress towards improvement Reducing variation identified as another interim marker consistent with a “get started” approach • Recognition of IHPs as an underutilized avenue for advancing • improvement 	<ul style="list-style-type: none"> • Clarify and more overtly track and celebrate measures of success with a “get started” approach

		<ul style="list-style-type: none">• Increased awareness of achievements worthy of celebration	
--	--	---	--

Chapter 9 Reflexivity in this action research project

Preamble

This chapter focusses on reflexivity in the action research project. The first section of the chapter outlines the purpose and nature of reflexivity involved in this action research. The second section summarizes my role in it as a scholar-practitioner and my reflections on the action research project as a whole. The focus of my reflexivity is based on the idea that “I am part of the problem and the problem is a part of me” (Laureate Online Education, 2010). The data sources for this exercise in reflexivity were notes to myself, notes to my academic supervisor and reflections during the process of writing the research thesis. I examined these data using the same reflective framework as the rest of the data in this action research project: what, so what and now what. The contribution of reflexivity to this action research is not only introspection to improve my own performance but more importantly insight into the factors contributing to progress on the problem addressed in this research.

Purpose of reflexivity

In action research, the researcher is inextricably involved in and affected by the research. The researcher is engaged in the work that is the subject of the research, trying to improve practice and at the same time study it to generate local and/or generalizable knowledge (Lyngsnes, 2016). While some would argue that researchers always affect and are affected by their research (Westerman, 2006), the connection between the researcher and the researched is intentionally and very obviously blurred in qualitative research in general, and insider action research in particular. Quality and rigour in this type of research depends on demonstrating that the researcher is aware of, and has taken into consideration, the extent to which they are

influencing the findings of the research (Sandelowski & Barroso, 2003; Ryan, n.d.). The search for and demonstration of this awareness is referred to as reflexivity.

If validity in qualitative research is related to the extent of behaviour change, as Bray et al., (2000) argue, reflexivity can increase the actual as well as the perceived validity of research. Reason (2006, p. 194) noted that there is a risk that the researcher, in their desire to be “helpful”, could push participants to act in ways that are not useful to them. Reflexivity can help the researcher see early signs of this and change their approach accordingly. The dialogue with participants that is part of the reflexive process can also increase the ownership of participants for the meaning and, more importantly, the implications for action emerging from the research (Lyngsnes, 2016). In this way, reflexivity increases the chances of meaningful action and thus the perceived validity and quality of the research.

Ethical behaviour in action research depends on reflexivity (Lyngsnes, 2016). An example of a reflexive stance in action research might be ongoing negotiation with research participants about the meaning of the observations being recorded (Doyle, 2007). The key word is “ongoing”. Lyngsnes (2016) observed that even when participants were not interested in what Doyle (2007) calls “participative member checking”, the very act of inviting it can build the trust between the researcher and the researched that is at the heart of ethical research.

Finally, the critical self-reflection that is part of reflexivity (Finlay & Gough, 2003, pp 3-20) is an important tool for ongoing professional development on the part of the researcher. However, care must be taken by the researcher to minimize the actual and perceived use of reflexivity simply for personal growth (Lyngsnes, 2016). Instead, researchers can use personal introspection to stimulate additional insights into their research. Lyngsnes (2016) heard from participants in her action research that “*Not any researcher would do however. It was very*

important that you once had been a teacher, that you know about new research and theory about classrooms, and that you are easy-going and speak in a way we understand!" In this way, the exploration into how participants perceived the researcher contributed to the knowledge generated by the research by suggesting specific attributes that are important for successful action researcher to embody (Lyngsnes, 2016).

The value of reflexivity is clear in the literature. Nevertheless, as Finlay (2002, p 209) notes, it is a messy process, "*full of muddy ambiguity and multiple trails as researchers negotiate the swamp of interminable deconstructions, self analysis and self disclosure*". To make it even more complex, there are numerous different versions of reflexivity described in the literature (Marcus, 1994; Wilkinson, 1988; and Finlay & Gough, 2003). To manage the "muddiness" of reflexivity in this action research, I address only 3 forms: reflexivity as introspection, intersubjective reflection and some degree of reflection as mutual collaboration. The application of these forms of reflexivity is described in more detail below.

Data sources for reflexivity

There were three data sources for this exercise in reflexivity. The first was notes to myself when I observed something surprising or otherwise interesting. My notes were not a journal per se as they were not recorded every day. The notes took the form of draft emails to myself and were stored together in a separate folder for the purpose. The second source of data was the compilation of notes addressed to my academic supervisor regarding my experience with doing and documenting the action research. Some of these notes were never sent as, on completion, they had served their purpose through the process of writing and sending them was therefore redundant. The final source of data was my reactions to the documentation of my observations. These are distinct from my reactions to the events in that they occurred much later

and were more focussed on how I had made meaning and how I had behaved at the time of the event being documented. For example, part of my reaction to low response rate on a survey was to decide against preparing a complete, formal summary of the results for distribution to members for (valid) reasons of low sample size and attendant issues with that. My reaction to the documentation of this reaction was a surprising insight about the potential that this decision might have contributed to continuing poor response or, at the very least, robbed the membership of the chance to reflect on their low response rate. Including this last source of data acknowledges the role of documentation of action research as an intervention or an “action” in its own right that deserves to be examined.

My role as scholar and practitioner

As described in Chapter 1, my roles in this action research project were as leader the initiative being investigated i.e. the implementation of D2D through the QIDS program of AFHTO and primary investigator. As such, my role is best described as an insider action researcher. I was not operating in a covert role as I had formal, explicit permission from the organization to use the D2D project as the subject of my doctoral research. In addition, the Board of AFHTO and steering committee for the QIDS program were reminded of the ongoing research process throughout the implementation of D2D. Evidence of their informed and active support include positive consideration of requests for tuition support and enrolment in research fellowships related to my program. Communication with members about my role as a researcher was more limited and showed up mostly in consent processes for specific activities like surveys or focus groups. However, all members were frequently reminded that I was actively examining “what works” with the program under the approval of Research Ethics Boards. The many presentations delivered in research forums were deliberately highlighted in communication with

members. Nonetheless, while not exactly “covert”, the daily activities of my research were intentionally casual and unheralded and intentionally rolled into my operational responsibilities. Not surprisingly, therefore, there was low awareness among members of the action research project. There was very high awareness of the actual initiative (i.e. D2D). People were also aware that I was intensely interested in feedback, that I was actively tracking it and that I was using the data being gathered to improve the initiative as well as to tell AFHTO’s story publicly and very positively in research forums and (with any luck) publications. There seemed to be lower awareness that I was always, in every interaction, in the business of data collection for my research. Evidence of this was the surprised congratulations from members when I would occasionally relay that I was in the process of finishing my thesis for my doctoral degree. On hearing that the topic of my research was “this” (that is, whatever we were doing or talking about), members would invariably be supportive and grateful that the work we were doing together was being taken so seriously. Occasionally, members with deep interests and experience in research pushed for (and were eventually satisfied with) more clarity about the potential conflict between my role as a researcher and my role to encourage members to participate in various activities. I remain immensely grateful for the trust placed in me by AFHTO and its members to deeply examine our experiment with performance measurement with an intent to help all of us get “even better than we were yesterday”. It has been a privilege to learn together with such inspiring classmates.

Reflections on my reflections

Apples and penguins: action learning and research

One idea that emerged as I reflected on my reflections was the difference between my experience of doing the earlier iterations of D2D and how I saw it through the lens of action

research. I expected that the action research project would be an opportunity to document what I had done and what I knew from my experience with D2D to date. I was hopeful that the action research process would be an opportunity to learn what might make D2D more successful. I was also hopeful that the process would help me grow as a scholar, particularly with respect to implementing methods consistent with my (albeit reluctant) constructivist philosophy. Because I had been deeply embedded in all aspects of D2D from the outset and had been very intentional about learning along the way, I did not truly expect to learn much about the implementation of D2D. That was my first surprise. Looking at D2D as an action research project showed me how little of the learning I had engaged in throughout D2D had been shared with anyone. Even though I actively sought and responded to member input from the very beginning, it was only after initiating the action research process that I became as deliberate and overt about referencing that. At first, it seemed patently ingratiating to add the phrase “in response to member input...” to my messages to members. It was only after doing it persistently for a period that I was able to see how much members were encouraged by this and how it further motivated ongoing engagement.

Another example was the reaction of a close and supportive partner in this work on hearing me speak for the first time about my view of D2D as an artefact being deployed according to activity theory: “so *THAT*’s what we’ve been doing all this time! *There really IS a method to the madness! Who knew?!*”. Prior to commencing this thesis, I thought action research was mostly a semantic distinction from action learning. I could be forgiven for not being really clear on exactly what action learning is since Revans himself (1981, p. 9) suggests that “the day action learning becomes explicable in words alone will be the day to abandon the practice of it”. Notwithstanding that, my experience in action research helped me see more

clearly the impact of sharing my reflections externally (via second person reflection or joint sense-making), a key distinction from my understanding of action learning. One member observed that comparing unlike teams to each other through D2D might result in not only comparing apples to oranges but even more dramatically, comparing “*apples to penguins*”. I am now seeing the apple/penguin divide between action learning and research. I can see how being more public in my reflections (as is necessary in action research) helps to translate internal learning on my part into knowledge that can help the community, which in turn can help me be more effective in my work on behalf of the community. I had looked forward with anticipation to the completion of this action research project and in particular, the end of my yearly reports to the research ethics board to maintain ethical approval for the research. I am now inclined to consider continued collaboration with the REB and other researchers to sustain a research focus in my operational work. My goal in this is not so much to continue to advance my research career but paradoxically to support my continued and growing effectiveness as a practitioner.

The fractal nature of D2D

Another far more uncomfortable surprise was that I was part of the problem I was researching. It was only on reflection on the low response rate from member surveys related to D2D that I could see my potential role in perpetrating that. I chose not to prepare or share results of surveys that had low response rates. I had what I thought then (and still think) were good reasons for those decisions. I just didn’t realize until reflecting on it that this choice was not consistent with the rest of the D2D philosophy. In choosing not to “get started” with whatever survey results I had, I was behaving in precisely the same way as the members I was trying to convince to just “do what you can” to take small steps forward. To be fair, I did use the data generated by the surveys. However, I did it almost covertly, thus depriving me and the members

of learning, not only about the value of the data, but also more about the apparent ambivalence between sharing performance but not team characteristic data. I also was surprised to see that I did not follow my own recommendations for action based on learning from the earlier iterations in D2D. Like the problem I perceived to be among the members, I did not act on what seemed clearly (at least to me) to be a valid and evidence-based thing to do. Again, I had (and have) reasons for that. They are not that interesting to me at the moment. Instead, the interesting part from the perspective of reflexivity is that I was exhibiting the same behaviour I was investigating on the part of the members.

This very uncomfortable realization is all the more embarrassing to me because it is not a new concept. The very first module of the DBA program introduced the concept that “I am part of the problem and the problem is a part of me”. In other words that are more aligned with my quantitative analytical background, my research is fractal: the same pattern of behaviour repeats at many levels. Using these terms makes it easier for me to understand and accept the idea by creating more distance between it and me. I don’t apologize for that. Learning is generally uncomfortable. However, too much discomfort can trigger defensive reactions and withdrawal. Finding ways to manage my discomfort and remain focussed on the problem (or rather, problems!) is not only my right but my responsibility in serving my organization. I am not immediately inclined to expose all my discomfort in this regard with my organization. As Morrison & Milliken (2000) observed, one person speaking out with deeply personal introspection is not only not likely to be effective in breaking organizational silence in this regard, but can also paradoxically perpetrate it. Instead, I am committed to finding ways to extend my inquiry into my own behaviour at least as deeply as my inquiry into others in a safe

way in my organization. Because many people have assured me that “*nobody reads a thesis*”, sharing these thoughts here feels like a safe place to start!

My story is, apparently, my story

Not all the surprises emerging from my reflections were as difficult. I was delighted to gradually gain confidence in and appreciation for qualitative methodology. This was an explicit personal goal for this action research. I had become increasingly dissatisfied with the capacity of quantitative methods (with which I have a high degree of comfort and confidence) to answer the questions that were emerging out of my work in performance measurement and quality improvement. I recognized a need for deeper and richer data about the experience of doing this work with those who I was trying to support. I was looking to build skills in qualitative methods to tell their story. I was particularly interested in finding tools that would convince me (and more importantly others) that the story I was telling was that of the organization and its partners, and not “my” story. I was thus grateful to have the opportunity and support to apply such a methodology (that is, template analysis as described by King (2004)) to help me do that. In fact, I became so comfortable in the method, I attempted to apply it to the reflexivity process for this thesis. I abandoned that when I realized that was exactly the opposite of what I had been trying to achieve in my quest for competence in qualitative methods. Instead of using them to get a richer understanding, I was trying to hide behind what had previously been foreign and unfamiliar to maintain a more comfortable distance between me and the story.

The experience of building skills in qualitative methods was satisfying not only for the confidence emerging from that but also from the realization that I do not have to remove myself from the story. It was a relief to release myself from the work of keeping my voice out of the story. I eventually realized that no methodology (qualitative or quantitative) could, or should,

take me out the story. I was the lead of the program that delivered the project that I initiated. The data was largely from conversations I had and meetings I attended. D2D can no more be separated from me than from those participating in it, or even those violently rejecting it. The emancipatory promise of criticality (Fournier & Grey 2000) had always appealed to me and I embraced it on behalf of those who I was serving. I was delighted to experience that it also applied to me. I was even more delighted to realize that being free to openly be part of the story was not just good for me but for the work itself. For example, in my efforts to increase collective ownership of the organization's measurement work, I intentionally use the first person plural pronoun. I am completely aware that this is sometimes more aspirational than accurate. While this *may* serve a purpose in my operational work, I came to learn that it was even more useful in my action research to be explicitly clear about my own actions, regardless of the nature of their impact. Being clear about my role in the problem creates more opportunities for me to examine my assumptions, thus enabling learning for me and on behalf of my organization. It also helps build the credibility of the story. When I can situate myself in all aspects of D2D, including those that did not work so well, I feel more confident in defending the credibility of the overall story to those with positivist inclinations (including me!) who are concerned about my bias.

Summary of reflections on my reflections

Reflexivity added another dimension to my action learning. Through reflection, I learned more about things I thought I already knew. I also learned that this improves not only the quality of my research, but also the quality of my practice. Reflection also helped me understand that I am in the problem and the problem is in me. I am now more focussed on finding ways to embrace this and at the same time manage how uncomfortable it is. Finally, reflecting on my

experience helped me see it more clearly and appreciate its value in improving myself *and* my work on behalf of my organization. It became clearer to me that reflexivity is not just navel gazing. The process did indeed help me improve myself but surprisingly to me, it also seemed applicable to the “keeping going” challenge we are facing with D2D. The challenge for me now is to find ways to build reflexivity into my approach to action learning to continually develop as an action researcher. I am not yet sure how well that will work. However, I am pretty sure that the best way to find out is to get started. After nearly 50,000 words elaborating ineffably on the concept of getting started, I think I should be ready to do it!

Appendix 1. Data dictionary

Cost	
DESCRIPTION	Indicator definition
	Per capita health care system cost with adjustment to reflect age/sex/complexity of patients.
	Reference Primary Care Performance Measurement Framework , see pg. 221. For more information, see Guidelines on Personal Level Costing
	Type Outcome Indicator
	External Alignment Primary Care Measurement Framework
DEFINITION & SOURCE INFORMATION	Unit of analysis Per capita
	Calculation See "Per capita health care expenditures by category" measure in the efficiency domain of the Primary Care Performance Measurement Framework , see pg. 221.
	Data source Primary Care Practice Team Report (ICES), see additional excel worksheet (addendum to core report): "Cost" Access via HQO Portal
	Data Elements <ul style="list-style-type: none"> • Total unadjusted Cost • Adjusted Total Cost • Primary Care Costs • Physician, Lab, drug, ED and outpatient Costs • Inpatient and same day surgery Costs • Long Term Care, Complex Continuing Care and Rehab Costs <p>*note: to be entered separately on data submission form. Please see PCPMF reference for descriptions of each cost element</p>
	Limitations/Caveats Some teams might not have access to the Primary Care Practice Team Report and therefore will not be able to report on this indicator.
OTHER RELEVANT INFORMATION	Rationale A measurement priority in the health system efficiency domain
	Drafted on Nov. 17, 2015
	Drafted by AFHTO Staff
	Updated on June 24, 2016
	Updated by AFHTO Staff
Update history	

Appendix 2. Coding template for action phase 1, Chapter 5

Aggregate Dimension	1 st level codes	2 nd level codes	Representative text
building relationships	characteristics in general	Agreeability	"Of course! I always do what Carol asks me to :)"
		Lighthearted exchanges	"Ah! You da best! (Sorry for the colloquialism with someone I don't know well, but when something is true, it just has to be said.)) "A new factor to investigate! Fun lol";
		Self deprecating exchanges	"My [local] solution not as refined as the [expensive alternative software] but I'm more of a beer guy than scotch anyways :)"
	conversations	AFHTO role as mediator	"I wondered if ministry funding will pay for family health team staff when they attend conference... Please kindly investigate"
		member input	"I wanted to provide all of you with some constructive feedback from our team ..."
		AFHTO offering help to external stakeholders	" I don't know how far along you are in your recruitment + selection process, but at this stage, would you like help from AFHTO? "
		AFHTO as a useful partner	"I was hoping we could include AFHTO as partners in the application"; "Thanks for your input on Monday - it has really helped to shape the grant"; "I think your info will speed the project along!"
		Appreciation of AFHTO input	"Thanks again for your comments. I've attached an updated version, which I believe addresses your concerns."
		Need for emotional support	"Whatever can be done to assist her would be appreciated. I would also appreciate it if you kept this to yourself as I know [she] would be embarrassed if she knew I shared this with you."
		building relationships	"Yes – it's getting legs! I finally have a connection with OMD at my level of interest and work – this is refreshing!"
		Introductions	"Dr S and Dr B are key leaders in our QI team, and you can certainly attend one of the monthly meetings to gain more on the ground feedback on D2D moving forwards"
		Leveraging relationships	"Hi friends. I hope you can help me send this out to your members."
		Appreciation of response to member input	"[members] are thrilled with the work that has been done advocating for the hospital discharge performance indicator to be changed. I was so excited to see the email come out last week with that information. "
		Interest in consensus	"great to see everyone on the working group on the same page!"
		Suspicious of government	"Make no mistake, AHFTO is an arm of the MOHLTC"; "another layer of bureaucracy in between the LHIN and the existing health groups.. Hilarious. What a dog's breakfast.";
		Low member engagement	"the rep for Executive Director Advisory Council in the area has not picked up the ball"
		seeking feedback	"soooooo, waddaya think????"; "Really appreciate you digging into this and sending us your feedback!"
	QI beliefs and definitions	Debates as a sign of engagement	"Interesting that anyone could be this belligerent about measurement. I didn't think that anyone really cared that much, so I take this as a good sign"
		Challenges in engaging physicians	"Given the toxic, bad-faith, heavy-handed relations with this government, I vote a strong NO to sharing data at this time... At this point, I don't even need to ""vote"": please do not share ANY of MY practice patient data."
		Resentment for poor performers	Advanced team feeling dragged down by poor performers

Aggregate Dimension	1 st level codes	2 nd level codes	Representative text
helplessness	characteristics in general	Solving own problems	"Going forward I would not leave the ""ask"" to the EDs ... but rather make a pitch to the Dr X and Dr Y myself a few months before the deadline."
		Help-seeking	"with regards to the QIDSS accessing FHT EMR's. Who is responsible for covering for the cost to the ERM vendor, the host FHT? "
		Permission seeking	"I will be asking 'permission' at tomorrow night's board meeting to request this data for the FHT"; "Can an FHT choose not to ask if there is little chance they could provide to patients should patient feedback indicate they want these offerings?"
		member inquiries	"is this an HQO expectation?"
		Technical problems	"I'm afraid the data you entered yesterday in the D2D Data Input page will need to be entered again"
		Scarce resources	"I am worried about the cost "
		Sharing information	"I created a board for us on Trello so we will all have access to up to date documents,"; "I discussed with FHT XX before sharing their data with you ... They were ok with that"
		Solutions from the field	"this is a great example of a team not receiving notifications in the EMR but how they have overcome this. MCFHT was able to be innovative and use resources and tools available to them to insure that they were able to provide discharge care whether or not an HRM/POI report was received."
		working together for more impact	"Thank you for sharing, I cannot speak for the rest of the QIDSS but this looks like it could evolve into a provincial tool."
		QIDSS helping QIDSS	"I can share the process map (swim lane diagram) later this week if you are interested."
	conversations	Difficulty finding data	"Because they are named differently, certain test results appear separately in the lab table, cannot be graphed, and cannot be searched on"
		Data capture	"The ehealth website talks about a new expedited process in getting registered. That's what I am interested in."
		Low EMR user knowledge	"A lot of users are just not aware ..."
		Physician-team division	"the "line" between FHT and FHO..."
		Diverse experiences with QIDSS	" their expectation of the QIDSS program was that they would have me 100% for four days a month working on solely [the local] FHT work. "; "our QIDSS is fantastic" "
	QI activities	QIDSS role in data management	"across the FHT, important data should be coded, should be standardized and structured, should be managed by the FHT and its Information Specialists, QIDSS. "
		Peer pressure drives data use	"a big part of [teams using D2D] is being able to highlight that they are being compared to similar teams."
low perceived priority	conversations	Carol talking a lot	"Any of you know this guy? He is a retired family phys and emerg doc ...interested in QI and change management etc -- met him at the... meeting today -- might be a good person to be on one of the working groups?" "that was part of the point of this exercise - ie to find out how resilient and resourceful they and the report functions are for getting at this and other data - so it's all good from a learning perspective"
		afhto internal discussions	"Here is the updated [AFHTO] communications plan – this what we'll all be working on over the next few weeks."; "ALL expense claims must be in by April 7 so that we can pay them out of this year's project funding".

Aggregate Dimension	1 st level codes	2 nd level codes	Representative text
qi activities	external help provided	External help provided	"I didn't expect to be coming back to you so quickly but we have the ability to export the data ...you are looking for and more importantly it is available right now."
		AFHTO as leader	"In our discussions with Ontario MOHLTC staff and Health Quality Ontario, AFHTO has been identified as a priority organization that we should include in our interview process."
		Not following AFHTO's lead	"There are no current plans [to follow AFHTO's lead in reporting]. We could raise it for discussion but it won't happen soon."
		Non-D2D communication with members	"we are sharing with you the response we received from each of the three main political parties "; "we have extended the deadline for Bright Lights nominations "
		communication about d2d	FYI – just sharing the good news that we launched the D2D 1.0 online tool today
		Connection of D2D to governance	"it was good to put be able to put D2D into the context of the new contract"
		Importance of stories	?? mostly carol text
	performance	Reluctance to share data showing value of teams	"proceed cautiously"
		Interest in D2D among members	"a few months ago most of my ED's are like.. D2D? Whats that? More work? yuck.... Now they are all like ""hey can we submit to D2D?"" after seeing it. ITS ALIVE!!!!!! Muahahahahhaaa!! Lol"
		D2D not an internal AFHTO priority	"If you think we need to move more quickly on Starfield story, let me know. I'm hoping the answer is that it can wait till end of March "
qi activities	characteristics in	Desires to acknowledge members	"Had a idea today. Wondered if we should put up a slide with all the working group members and the advisory panel"
		Interest in celebration	"we can reflect with pride on the quality of care provided to patients"
		Focus on getting started	"We'll see how the first one goes"; " I think messiness may be worth trying"
		collaborating	We pride ourselves in working collaboratively with other organizations to support our members
	collaboration	Lack of provider interest in patient engagement	"I did check with 3 of my FHTs but none of them interested [in inviting a patient to attend a joint IHP/QIDSS/patient session]"
		Patient experience surveys as obligation	"so you have always done your survey in December – does it ALWAYS have to be that way going forward if there is a compelling reason not to?"
		Patient input	" I really appreciated being able to table some of my issues as a patient... The perogies were amazing, too!"
		Example of presenting data	"No fancy graphs right now, but I do have a very colourful ... spreadsheet!"
		Examples of using data	"knowing that we are not doing as well as we think we should is an important part of deciding to do better"
		Privacy as barrier to participation	"there are some concerns/hesitations from physicians regarding confidentiality of patient information thus I have not been able to get the sign off to participate".

Appendix 3. Coding template for action phase 2 (experience with D2D 4.0), Chapter 7

Aggregate dimension	1st level codes (from 1st cycle)	2nd level codes (from data)	Representative text
Getting started	Celebrate	actual celebrations	"I'll be getting the cake for her real celebration next month
		compliment	This is SUCH great stuff!!
		stuff to celebrate	Deliver good news messaging about D2Dd to internal and external stakeholders. There are some real success stories coming out of the work the QIDS program is doing
		stories vs success stories	Many fhts [feel] it would seem presumptuous to claim what a fht is doing qualifies as a "success" If there were a simpler invitation to describe programs that fhts are pleased about, people might feel more comfortable to contribute.
	Get started at keeping going	approach	" Biggest priority for AFHTO is to promote the approach
		good enough now	To be expected for first few years/iterations – might be more reasonable to expect improvement by 2018-19 Some stakeholders were surprised that it was an expectation that measurement as intended to lead to improvement,
		measure measurement	A continued focus on improving the quality of measurement is also important to make sure we continue to consider: "are we measuring the right things?"
		program planning indicators	The Schedule A program planning guides and resources (including the indicator catalogue) are intended to help teams demonstrate value for the money invested in their programs.
		reduce variation	Identify teams who not only were the best performers but teams that also showed the greatest improvement – this will help other teams decide who they want to talk to help them improve.
		tools	"Members agreed with the concept of introducing an "improvement" initiative to build off the "measurement" initiative (ie D2D)
		use ihsps	IHPs are involved in... Schedule A with very little or no involvement from physicians. IHP have a vested interest in setting the indicators for Schedule A because they are usually asked to contribute data and also the data reflect progress on the programs they lead
		demonstration of increased use	Define indicators of "use" such as: identify action to be taken in response to the data and identify specific patients to take action upon [and several other indicators]
		using data	use report to stimulate conversation about improving quality of, access to & use of EMR data
		facilitate use	key influencers/enablers or motivators to leverage in developing the information package to increase use of D2D to improve were [list]
	Stay small	make it easier	Having a place to highlight and celebration innovation (Eg innovation symposium of may 2014) would provide a great opportunity to spread message
		alignment with requirements	Continue to educate members about the alignment/connection between D2D and other reporting processes eg schedules A and E and QIP
		feedback mechanism	there is a gap in the feedback mechanism in D2D.
		set expectations	Incorporate thresholds from member survey as "targets"
		stay small	Keep the number of indicators PRESENTED in D2D small (12-20).
	stage-specific interventions	contemp	Teams need to be nurtured in this as it is difficult work
		stage-specific interventions	Members must be able to see the value in participating is more than the [data capture] burden

Aggregate dimension	1st level codes (from 1st cycle)	2nd level codes (from data)	Representative text
Help seeking/self reliance	Consider stages of change	precontemp	Large gaps in reporting could potentially diminish the momentum of the project and diminish member knowledge. High frequency is important in QI – preferred frequency is quarterly but this is not feasible for D2D.,
		prep	The HQO portal providing EDs with access to team-level [reports] has now opened
Perceived priority of D2D	Communicate the value of D2D more effectively	develop more content	better data (ie from validation study) would strengthen ability to advocate for AFHTO members in discussions with MOHLTC/LHINs
		examples of success in communication	Members are pleased with the introduction of the printed report [ie in addition to the interactive version on the web site]; The ability to see data in trend format considered valuable; Teams liked the new Quality roll-up drill down
		messaging suggestions	greater emphasis ... on D2D's influence on MOHLTC, how data is of value to team
		process suggestions	Members also identified the need for a cheat sheet to help with one-on-one conversations
	Explore theories in use	not there yet	Lack of awareness of purpose [of D2D]; No sense of urgency in doing something [to improve]
relationships	Build cadre of physician champions	improvement in engagement	Team tracks who reviews the website and see that 2-3 physician champions looking at it a lot.
		little physician engagement	Had to hunt down the necessary clinicians to help complete the survey.
		strategies to improve engagement	Sharing D2D and physician specific data online takes advantage of competitive nature of docs to stimulate improvement.
		incorporating member input	Based upon the physician input, it was deemed that is would not be possible...
	Continue inviting participation	build phys champs	Send out personal message to medical leads
		example of input	Consider reducing number of surveys (note the need to balance with the need to get member input especially for D2D to evolve as desired by members – ie be more meaningful).
	Maintain attention on conversations	impact of input	I have to say that I was delighted with the process as we are trying to keep D2D "meaningful" and the providers' voices were listened to.
		seek input	Get feedback on video from EDs on what would be useful to them
	Nurture relationships	inviting participation	Focus energies on teams with no QIDSS support
		nurture relationships	Wording should be sensitive to avoid perceptions of judgment or blame
		building relationships	this will both be informative but also a great networking opportunity for our project.
		leverage relationships	AFHTO to approach high performers for stories
		protecting relationships	Because of [the risk in breaching the precious trust that teams have in AFHTO], IWG members recently decided AGAINST incorporating D2D into QIP and PCR reports, even tho this would meet members' needs for ... reduced duplication of reporting efforts.
other	other	changed indicators in D2D	teams will also have the option to indicate their LHIN [geographical area]
		make D2D easier to use	Members agreed that the multiple sign-ups (for d2d 2.0 and for HQO and then for D2D 3.0) were confusing and not necessary since the D2D signups are not actually required. Therefore IWG decided to eliminate the sign-up process for subsequent iterations of D2D.

Appendix 4. Abstracts submitted and accepted for presentation at various conferences

Nine abstracts based on various aspects of the first three iterations of D2D were submitted to various conferences. All nine were accepted for presentation as noted below:

Trillium Primary Health Care Research Day, Toronto, Ontario, June 1, 2016

1. Mulder, C., Leyland, M. & Asalya, M. (2016) *Feasibility and Impact of Using EMR to Trigger Automated Patient Experience Surveying*, Available at:
http://www.trilliumresearchday.com/documents/2016_201D_Mulder_Asalya_Feasibility_and_impact_of_using_EMR_to_trigger_1.pdf (Accessed Nov 5, 2017).
2. Mulder, C., Zago, D. & Wilkerson, T. (2016) *Getting Started with Involving Patients in Improving Quality*, Available at:
http://www.trilliumresearchday.com/documents/2016_203D_Mulder_Zago_Getting_start_ed_with_involvin_g_patients_in_improving_quality.pdf (Accessed Nov 5, 2017).
3. Mulder, C., Glazier, R. & Sullivan, F. (2016) *Ontario Data Support Starfield's Theory on Primary Care Quality and Cost*, Available at:
http://www.trilliumresearchday.com/documents/2016_203D_Mulder_Ontario_data_support_Starfields_theory_on_primary_care.pdf (Accessed Nov 5, 2017).
4. Mulder, C. & Wilkerson, T. (2016) *What do Interprofessional Health Care Providers Need and Want to Get Better at What They do?*, Available at:
http://www.trilliumresearchday.com/documents/2016_203D_Mulder_Wilkerson_What_d_o_interprofession al_healthcare_providers_need_and_want.pdf (Accessed Nov 5, 2017).

North American Primary Care Research Group Practice-Based Research Conference,
Bethesda, USA, July 11–12, 2016

5. Mulder, C., Sullivan, F. & Greenberg, A. (2016) *Making composite measures of quality useful for front-line primary care providers*, (pp 13-14, code: WPF10), Available at:
<http://www.napcrg.org/Portals/51/Documents/PBRN%20Meeting/2016%20Meeting/Wor kshop%20Panel%20Forum%20Abstracts%202016.pdf?ver=2016-08-03-081110-057> (Accessed Nov 5, 2017).
6. Mulder, C. & Glazier, R. (2016) *Impact of a ground-up voluntary performance measurement initiative on the use of data for QI in primary care*, (pg. 10, code: OP10)
Available at:
<http://www.napcrg.org/Portals/51/Documents/PBRN%20Meeting/2016%20Meeting/Abstracts%202016%20OP%20FINAL.pdf?ver=2016-08-03-080525-080> (Accessed Nov 5, 2017).
7. Mulder, C., Glazier, R., Sullivan, F. & Southey, G. (2016) *Ontario Data Support Starfield's Theory on Practice Quality and Cost*, (pg. 20, code: OP19), Available at:
<http://www.napcrg.org/Portals/51/Documents/PBRN%20Meeting/2016%20Meeting/Abstracts%202016%20OP%20FINAL.pdf?ver=2016-08-03-080525-080> (Accessed Nov 5, 2017).

Ontario College of Family Physicians 54th Annual Scientific Assembly, Toronto, Ontario,

November 24-26, 2016

8. Mulder, C. (2016) *Reduce, Reuse, Recycle: Digging for Gold in EMR Data*, Available at:
<https://www.eventscribe.com/2016/asa/aasearchbyspeaker.asp?h=Browse by Speaker>
(Accessed Nov 5, 2017).

Canadian College of Family Physicians Family Medicine Forum, Vancouver, British Columbia,
November 9-12, 2016

9. Mulder, C., Glazier, R., Sullivan, F. & Southey, G. (2016) *Ontario Data Support Starfield's Theory on Practice Quality and Cost*, (Session code: W141717) Available at:

https://fmf.cfpc.ca/wp-content/uploads/2016/12/FMF_2016_Final_Program.pdf

(Accessed Nov 5,

2017).

References

- ‘A Christmas Story’ (1983) Wikipedia. Available at:
https://en.wikipedia.org/wiki/A_Christmas_Story (Accessed Sep 2, 2017).
- Aggarwal, M. (2009) Primary Care Reform: A Case Study of Ontario, PhD Thesis. University of Toronto.
- Aguinaldo, J.P. (2004) ‘Rethinking validity in qualitative research from a social constructionist perspective: From “is this valid research?” to “what is this research valid for?”’, *The Qualitative Report*, 9(1), pp. 127-136.
- Ajzen, I. (1991) ‘The theory of planned behavior’, *Organizational Behavior and Human Decision Processes*, 50 (2), pp. 179–211.
- Allen, M., Ferrier, S., O’Connor, N. & Fleming, I. (2007) ‘Family physicians’ perceptions of academic detailing: a quantitative and qualitative study’, *BMC Medical Education*, 7 (36).
- Anderson, C.J. (2003) ‘The Psychology of Doing Nothing: Forms of Decision Avoidance Result From Reason and Emotion’, *Psychological Bulletin*, 129(1), pp. 139–167.
- Argyris, C. (1996) ‘Unrecognized Defenses of Scholars: Impact on Theory and Research’, *Organization Science*, 7 (1), pp. 79-87.
- Argyris, C. (1999) *On Organizational Learning*. 2nd ed. Blackwell Business.
- Argyris, C. & Schon, D. (1974) *Theory in practice: Increasing professional effectiveness*. Jossey Bass.
- Argyris, C. & Schon, D. (1978) *Organisational learning: A theory of action perspective*. Addison Wesley

- Armenakis, A.A. & Bedeian, A.G. (1999) 'Organizational change: a review of theory and research in the 1990s', *Journal of Management*, 25 (3), pp. 293–315.
- Aronson, E. (1992) 'The Return of the Repressed: Dissonance Theory Makes a Comeback', *Psychological Inquiry*, 4, p. 303.
- Arrata, P. (2007) 'Building an effective change agent team', *The McKinsey quarterly*, 4, pp. 39-43.
- Associates in Process Improvement (API), (2016) *Model for improvement*. Available at:
<http://www.apiweb.org.> (Accessed Jan 31, 2016).
- Association of Family Health Teams of Ontario (2012) *Toward a Primary Care Recruitment & Retention Strategy for Ontario*. Available at:
<http://www.afhto.ca/highlights/news/toward-a-primary-care-recruitment-retentionhttp://www.afhto.ca/highlights/news/toward-a-primary-care-recruitment-retention-strategy-for-ontario-2/strategy-for-ontario-2/> (Accessed March 2016).
- Association of Family Health Teams of Ontario (2013) *Strategic Direction*. Available at:
<http://www.afhto.ca/wp-content/uploads/AFHTO-Strategic-Direction-2013-03-06.pdf>
(Accessed Jun 12, 2017).
- Association of Family Health Teams of Ontario (2016) *Quality Roll-Up Indicator: The Movie!*. Available at:
https://www.youtube.com/watch?v=CxGUZaFMPjU&feature=youtu.be&list=PLN8BxMxtC-XXApGFktOWazOA3QT5mt_NL (Accessed Jun 12, 2017).
- Bahamon, C., Dwyer, J. & Buxbaum, A. (2006) 'Leading a change process to improve health service delivery', *Bulletin of the World Health Organization*, 84(8), pp. 658-661.

- Baily, M., Bottrell, M., Lynn, J. & Jennings, B. (2006) 'The ethics of using QI methods to improve health care quality and safety', *Hastings Center Report*, 36(4), pp. S1-40.
- BakerGoodChild (2016) *Advantages and Disadvantages of Direct Marketing*, Available at: <https://www.bakergoodchild.co.uk/advantages-and-disadvantages-of-direct-marketing/>, (Accessed Sep 14, 2017).
- Baptiste, N.R. (2009) 'Fun and well-being: insights from senior managers in a local authority', *Employee Relations*, 31(6), pp. 600-612.
- Bazerman, M.H., & Samuelson, W.F. (1983) 'I Won the Auction but Don't Want the Prize', *The Journal of Conflict Resolution*, 27(4), pp. 618-634.
- Bechky, B.A. (2003) 'Object Lessons: Workplace Artifacts as Representations of Occupational Jurisdiction', *American Journal of Sociology*, 109(3), pp. 720-752.
- Berends, H., Boersma, K. & Weggeman, M. (2003) 'The Structuration of Organizational Learning', *Human Relations* 56(9), pp. 1035-1056.
- Björkman, H. & Sundgren, M. (2005) 'Political entrepreneurship in action research: learning from two cases', *Journal of Organizational Change Management*, 18 (5), pp.399-415.
- Blackler, F. (2011) 'Power, politics, and intervention theory: Lessons from organization studies' *Theory & Psychology* 21(5), pp. 724–734.
- Bowen, M. (1987) 'The Escalation Phenomenon Reconsidered: Decision Dilemmas or Decision Errors?', *Academy of Management Review*, 12(I), pp. 52-66.
- Bowen, M. & Lau, F. (2012) 'Defining and Evaluating Electronic Medical Record Data Quality within the Canadian Context', *Electronic Healthcare*, 11(1), e5-e13.
- Bray, J. N., Lee, J., Smith, L. L. & Yorks, L. (2000) *Collaborative inquiry in practice: Action, reflection and making meaning*. Sage.

Brockner, J. & James, E.H. (2008) 'Toward an understanding of when executives see crisis as opportunity', *Journal of Applied Behavioral Science*, 44 (1), pp. 94–115.

Brown, J. (2010) 'Do You Have Physician Support and a Physician Champion at Your Side?', *Journal of Health Care Compliance*, 12(6), pp. 39-40.

Burnes, B. (2004) 'Kurt Lewin and the Planned Approach to Change: A Re-appraisal', *Journal of Management Studies*, 41 (6) pp. 977-1002.

Canada. Canadian Environmental Assessment Agency (2016), *Public Participation Guide: Chapter 5. Implementing Public Participation*, Available at:

<https://www.canada.ca/en/environmental-assessment-agency/services/policyhttps://www.canada.ca/en/environmental-assessment-agency/services/policy-guidance/public-participation-guide/implementing-public-participation.html - p1-6guidance/public-participation-guide/implementing-public-participation.html#p1-6>, (Accessed Sep 13, 2017).

Canadian Foundation for Health Improvement (CFHI) (2013) *Patient engagement: heard and valued*. Available at: http://www.cfhi-fcass.ca/sf-docs/default-source/patienthttp://www.cfhi-fcass.ca/sf-docs/default-source/patient-engagement/awesome_handbook-fraserhealth.pdf?sfvrsn=2 (Accessed May 20, 2017).

Canadian Foundation for Health Improvement (2014) *Making Data Matter: Measurement Basics for Quality Improvement*. Available at: <http://www.cfhi-fcass.ca/WhatWeDo/onhttp://www.cfhi-fcass.ca/WhatWeDo/on-call/measurement-for-quality-improvement-webinar-series/making-data-matter-measurement-basics-for->

[quality-improvementcall/measurement-for-quality-improvement-webinar-series/making-data-matterhttp://www.cfhi-fcass.ca/WhatWeDo/on-call/measurement-for-quality-improvement-webinar-series/making-data-matter-measurement-basics-for-quality-improvementmeasurement-basics-for-quality-improvement](http://www.cfhi-fcass.ca/WhatWeDo/on-call/measurement-for-quality-improvement-webinar-series/making-data-matterhttp://www.cfhi-fcass.ca/WhatWeDo/on-call/measurement-for-quality-improvement-webinar-series/making-data-matter-measurement-basics-for-quality-improvementmeasurement-basics-for-quality-improvement), (Accessed May 20, 2017).

Canadian Institute for Health Information (CIHI) (2016). *How Canada Compares: Results From The Commonwealth Fund 2015 International Health Policy Survey of Primary Care Physicians*. Available at:

https://www.cihi.ca/sites/default/files/document/commonwealth_fund_2015_pdf_en.pdf
(Accessed Jan 7, 2016).

Canadian Medical Association (CMA) (2014) *How can Canada achieve enhanced use of electronic medical records?* Available at:

<https://www.cma.ca/Assets/assetshttps://www.cma.ca/Assets/assets-library/document/en/advocacy/Enhanced-Use-of-EMRs-Discussion-Paper-Final-May-2014.pdflibrary/document/en/advocacy/Enhanced-Use-of-EMRs-Discussion-Paper-Final-Mayhttps://www.cma.ca/Assets/assets-library/document/en/advocacy/Enhanced-Use-of-EMRs-Discussion-Paper-Final-May-2014.pdf2014.pdf> (Accessed Sep, 14, 2017).

Capelli, T. & Smithies, A. (2009) ‘Online Communities of Practice as Agents of Change in Curriculum Development’ *International Conference on e-Learning*, Toronto, Canada Jul 16-17. Available at: <https://www.escholar.manchester.ac.uk/uk-ac-man-scw:78654>
(Accessed May 15, 2016).

Carmeli, A. & Schaubroeck, J. (2008) ‘Organisational crisis-preparedness: The importance of learning from failures’, *Long Range Planning: International Journal of Strategic Management*, 41 (2), pp. 177–196.

- Carman, K.L., Dardess, P., Maurer, M., Sofaer, S., Adams, K., Bechtel, C. & Sweeney, J. (2013) ‘Patient and family engagement: A framework for understand the elements and developing intervention and policies’, *Health Affairs*, 32(2), pp. 223-231.
- Cassel, C.K., Conway, P.H., Delbanco, S.F., Jha, A.K., Saunders, R.S. & Lee, T.H. (2014) ‘Getting More Performance from Performance Measurement’, *N Engl J Med* 371, pp. 2145-2147.
- Cennamo, C., Berrone, P. & Gomez-Mejia, L.R. (2009) ‘Does stakeholder management have a dark side?’, *Journal of Business Ethics*, 89(4), pp. 491-507.
- Chugh, D. & Bazerman, M. (2007) ‘Bounded awareness: what you fail to see can hurt you’, *Mind & Society*, 6, pp. 1–18.
- Churchman, C.W. (1967) ‘Wicked problems’, *Management Science*, 14 (4), pp.141-42.
- Coghlan, D. & Brannick, T. (2014) *Doing action research in your own organization*. 4th ed. Sage.
- Cook, T. (2006) ‘Collaborative Action Research within Developmental Evaluation: Learning to See or the Road to Myopia?’, *Evaluation*, 12(4), pp.418– 436.
- Crabtree, B.F., Miller, W.L., Tallia, A.F., Cohen, D.F., DiCicco-Bloom, B., McIlvain, H.E., Aita, V.A., Scott, J.G., Gregory, P.B., Stange, K.C. & McDaniel, R.R.Jr. (2005) ‘Delivery of Clinical Preventive Services in Family Medicine Offices’, *The Annals of Family Medicine*, 3, pp.430-435.
- Creswell, J. (2007) *Qualitative inquiry and research design: choosing among five approaches*. 2nd ed. Sage.
- Davis, D., Mazmanian, P.E., Fordis, M., Van Harrison, R., Thorpe, K.E. & Perrier, L. (2006) ‘Accuracy of Physician Self-assessment Compared With Observed Measures of

Competence: A Systematic Review', *Journal of the American Medical Association*, 296(9), pp.1094-102.

de Lusignan, S. & van Weel, C. (2006) 'The use of routinely collected computer data for research in primary care: opportunities and challenges', *Family Practice*, 23, pp. 253–263.

De Silva, D. & Bamber, J. (2014) *Improving quality in general practice*. Available at: <http://www.health.org.uk/publication/improving-quality-general-practice> (Accessed Jul 12, 2017).

de Vries, S., Voorham, J., Haaijer-Ruskamp, F., & Denig, P. (2014) 'Potential Over-treatment and Under-treatment of Diabetes in Different Patient Age Groups in Primary Care After the Introduction of Performance Measures', *Diabetes Care*, 37(5), pp. 1312-1320.

Deken, F., Carlile, P.R., Berends, H. & Lauche, K. (2016) 'Generating Novelty Through Interdependent Routines: A Process Model of Routine Work', *Organization Science*, 27(3), pp. 659-677.

Derrick-Mills, T. (2015) *Understanding Data Use for Continuous Quality Improvement in Head Start: Preliminary Findings*, Office of Planning, Research and Evaluation Report # 2015-33. Available at:

https://www.acf.hhs.gov/sites/default/files/opre/brief_cqi_in_head_start_finaldraftcleanv2_508.pdf (Accessed Apr 18, 2016).

Dervitsiotis, K.N. (2002) 'The importance of conversations-for-action for effective strategic management', *Total Quality Management*, 113(8), pp. 1087-1098.

Dick, B. (2009) 'Action research literature 2006–2008: Themes and trends', *Action Research*, 7(4), pp. 423–441.

Dooley, K.J. (1997) 'A complex adaptive systems model of organizational change', *Nonlinear Dynamics, Psychology and Life Science*, 1 (1), pp. 69–97.

Dorazio, J. (2014) *Citizen engagement: 4 key insights from data on citizen satisfaction*. Available at: <https://www.visioncritical.com/citizen-engagement-4-key-insights-data-citizen-satisfaction/satisfaction/>, (Accessed Sep 13, 2017).

Doyle, S. (2007) 'Member Checking With Older Women: A Framework for Negotiating Meaning', *Health Care for Women International*, 28(10), pp. 888-908.

Drucker, P. (1973) *Management: Tasks, Responsibilities and Practices*. William Heinemann Ltd.

Edmondson, A. (2003) 'Speaking Up in the Operating Room: How Team Leaders Promote Learning in Interdisciplinary Action Teams', *Journal of Management Studies*, 40(6), pp. 1419-1452.

Edmondson, A., Bohmer, R.M. & Pisano, G.P. (2001) 'Disrupted Routines: Team Learning and New Technology Implementation in Hospitals', *Administrative Science Quarterly*, 46, pp. 685-716.

eHealthOntario (2015) *Benefits Realization Update 2015*. Available at: https://www.ehealthontario.on.ca/images/uploads/annual_reports/Benefits_Realization_Update-en.pdf (Accessed Nov 20, 2016).

Engeström, Y. (2000) 'Activity Theory as a Framework for Analyzing and Redesigning Work', *Ergonomics*, 43(7), pp. 960-74.

Ezzamel, M., Willmott, H. & Worthington, F. (2001) 'Power, control, and resistance in the factory that time forgot', *Journal of Management Studies*, 38 (8), pp. 1053–1079.

Family Medicine for America's Health (n.d.). Available at: <http://fmahealth.org/about> (Accessed

No6, 2016).

Finlay, L. (2002). 'Negotiating the swamp: the opportunity and challenge of reflexivity in research practice', *Qualitative Research*, 2(2), pp. 209-230.

Finlay, L. & Gough, B. (2003) *Reflexivity. A practical guide for researchers in health and social sciences*. Blackwell Publishing.

Fook, J. & Askeland, G.A. (2007) 'Challenges of Critical Reflection: Nothing Ventured, Nothing Gained', *Social Work Education*, 26(5), pp. 520–533.

Ford, J.D. & Ford, L.W. (1994) 'Logics of identity, contradiction and attraction in change', *Academy of Management Review*, 19 (4), pp. 756–85.

Ford, J.D. & Ford, L.W. (1995) 'The role of conversations in producing intentional change in organizations', *Academy of Management Review*, 20(3) pp. 541-570.

Ford, J.D., Ford, L.W. & D'amelio, A. (2008) 'Resistance to change: The rest of the story', *Academy of Management Review*, 33 (2), pp. 362–377.

Fournier, V. & Grey, C. (2000) 'At the critical moment: conditions and prospects for critical management studies', *Human Relations*, 53 (1), pp.7-32.

Gabriel, R.P. (1992), 'Is worse really better?', *Journal Of Object Oriented Programming*, 6, p. 3.

Gallagher, K., Nutting, P.A., Nease, D.E.Jr., Graham, D.G., Bonham, A.J., Dickinson, W.P. & Main, D.S. (2010) 'It Takes Two: Using Coleaders to Champion Improvements in Small Primary Care Practices', *The Journal of the American Board of Family Medicine*, 23(5), pp. 632-639.

Gamble, J.A.A. (2008) *A Developmental Evaluation Primer*. The J.W. McConnell Family Foundation.

Giddens, A. (1984) *The Constitution of Society*. University of California Press.

- Gioia, D.A., & Chittipeddi, K. (1991) ‘Sense-making and sense-giving in strategic change initiation’, *Strategic Management Journal*, 12 (6), pp. 433–448.
- Grant, K. (2015) ‘Ontario’s curious shift away from family health teams’, *Globe and Mail*, Feb 15. Available at: <http://www.theglobeandmail.com/life/health-and-fitness/health/ontarioshttp://www.theglobeandmail.com/life/health-and-fitness/health/ontarios-curious-shift-away-from-family-health-teams/article22989363/> (Accessed May, 12, 2017).
- Green, M.E., Khan, S., Frymire, E., Kopp, A., Kiran, T. & Glazier, R.H. (2016) ‘Association Between Patient Perceptions of Access to Care and Emergency Department Use-2013-2015’. *NAPCRG Annual Conference*. Nov 12-16. Available at: <http://www.napcrg.org/Conferences/AnnualMeeting/SearchEducationalSessions-2016?language=en-US&m=2&s=Association%2520Between%2520Patient%2520Perceptions%2520of%2520Access%2520to%2520Care%2520and%2520Emergency%2520Department%2520Use&p=0> (Accessed Nov 20, 2016).
- Greenwood, D.J. & Levin, M. (2007) *Introduction to action research*. 2nd ed. Sage.
- Greiver, M., Barnsley, J., Glazier, R.H., Moineddin, R. & Harvey, B.J. (2011) ‘Implementation of electronic medical records: Effect on the provision of preventive services in a pay-for-performance environment’, *Canadian Family Physician*, 57, pp. e381-e389.
- Greiver, M., Aliarzadeh, B., Williamson, T., Barber, D., Birtwhistle, R., Khan, S., Morkem, R.,

- Halas, G., Katz, A. & Harris, S. (2014) 'Prevalence and Epidemiology of Diabetes in Canadian Primary Care Practices: A Report from the Canadian Primary Care Sentinel Surveillance Network' *Canadian Journal of Diabetes*, 38(3), pp. 179-185.
- Grint, K. (2005) 'Problems, problems, problems: The social construction of 'leadership'', *Human Relations*, 58 (11), pp. 1467-1494.
- Guthrie, B. (2008) 'Measuring the quality of healthcare systems using composites', *British Medical Journal*, 337, pp 7667.
- Hall, F.M. (2014) 'Screening mammography guidelines: an alternative proactive approach', *Radiology*, 273(3), pp. 646-651.
- Health Quality Ontario (HQO) (2013) *Measurement for Quality Improvement*. Available at: <http://www.hqontario.ca/Portals/0/Documents/qi/qi-measurement-primer-en.pdf> (Accessed Feb 2, 2017).
- Health Quality Ontario (HQO) (2014) *A Primary Care Performance Measurement Framework for Ontario: Report of the Steering Committee for the Ontario Primary Care Performance Measurement Initiative: Phase One*. Available from: <http://www.hqontario.ca/Portals/0/documents/pr/pc-performance-measurement-report.pdf> (Accessed Apr 12, 2015).
- Heller, F. (2004) 'Action Research and Research Action: A Family of Methods', in Cassell, C. & Symon, G. Essential Guide to Qualitative Methods in Organizational Research, Sage, pp. 256-270.
- Hospital Research and Education Trust (1992), *Bridging the gap between theory and practice:*

Exploring continuous quality improvement. Available at:

<https://intermountainhealthcare.org/-/media/files/research/healthcare-delivery-articles>
<https://intermountainhealthcare.org/-/media/files/research/healthcare-delivery-articles-good-enough.pdf?la=eng&good-enough.pdf?la=en>, (Accessed May 20, 2017).

Howell, M.T. (2006) *Actionable Performance Measurement: A key to success*. ASQ Quality Press.

Hutchison, B. & Glazier, R. (2013) ‘Ontario’s Primary Care reforms have transformed the local care landscape, but a plan is needed for ongoing improvement’ *Health Affairs*, 32(4) pp. 695–703.

Institute for Healthcare Improvement (IHI) (n.d.) *When Good Enough Isn’t.... Good Enough: The Case for Reliability*. Available at:

<http://www.ihi.org/resources/Pages/ImprovementStories/WhenGoodEnoughIsn'tGoodEnoughTheCaseforReliability.aspx> (Accessed Aug 12, 2017).

Institute for Healthcare Improvement (IHI), (2016) *Science of improvement*. Available at:

<http://www.ihi.org/about/Pages/ScienceofImprovement.aspx> (Accessed Aug 24, 2016).

Ivers, N.M., Sales, A., Colquhoun, H., Michie, S., Foy, R., Francis, J. J. & Grimshaw, J.M. (2014) ‘No more ‘business as usual’ with audit and feedback interventions: towards an agenda for a reinvigorated intervention’, *Implementation Science*, 9(14).

Janis, I.L. (1973) ‘Groupthink and group dynamics: A social psychological analysis of defective policy decisions’, *Policy Studies Journal*, 2(1), pp. 19-25.

Janney, J.J. & Dess, G.G. (2004) ‘Can real-options analysis improve decision-making? Promises and pitfalls’, *Academy of Management Executive*, 18(4), pp. 60-75.

Jenkins, T. (2016) *What are the top challenges facing membership organisations?* Available at:

<http://www.pagelizard.com/top-challenges-facing-membership-organisations/> (Accessed Sep 3, 2017).

Johnson, P. & Duberley, J. (2003) ‘Reflexivity in management research’, *Journal of Management Studies*, 40 (5), pp. 1279–1303.

Kajamaa, A. (2012) ‘Enriching action research with the narrative approach and activity theory: analyzing the consequences of an intervention in a public sector hospital in Finland’, *Educational Action Research*, 20(1), pp. 75–93.

Kanouse, D.E., Kallich, J.D., Kahan, J.P. (1995) ‘Dissemination of effectiveness and outcomes research’, *Health Policy* 34, pp. 167-192.

Kaplan, S.H., Griffith, J.L., Price, L.L., Pawlson, L. G., Greenfield, S. (2009) ‘Improving the Reliability of Physician Performance Assessment: Identifying the ‘Physician Effect’ on Quality and Creating Composite Measures’, *Medical Care*, 47(4), pp. 378-387.

Kass, N. & Pronovost, P. (2011) ‘Quality, safety, and institutional review boards: navigating ethics and oversight in applied health systems research’, *American Journal of Medical Quality*, 26(2), pp. 157-159.

Keshavjee, K. (2007) *EMR implementation in Ontario: Position paper to increase the deployment of Electronic Medical Records in Ontario*. Available at: http://blog.canadianemr.ca/Articles/Intel_InfoClin_EMR_White_Paper_Final_080125.pdf (Accessed: Mar 14, 2016).

Kemmis, S. & McTaggart, R. (eds) (1988) *The Action Research Planner*. Deakin University Press.

Kilduff, M. & Dougherty, D. (2000) ‘Change and development in a pluralistic world: the view from the classics’, *Academy of Management Review*, 25 (4), pp. 777–782.

- Kim, L. (1998) ‘Crisis Construction and Organizational Learning: Capability Building Catching-up at Hyundai Motor’, *Organization Science*, 9(4), pp. 506-521.
- King, N. (2004) ‘Using Templates in the Thematic Analysis of Text’, in Cassell, C. & Symon, G. *Essential Guide to Qualitative Methods in Organizational Research*, Sage, pp. 256270.
- Kiran, T. & Obrien, P. (2015) ‘Challenge of same-day access in primary care’, *Canadian Family Physician*, 61, pp. 399-400.
- Kirchner, J.E., Parker, L.E., Bonner, L.M., Fickel, J.J., Yano, E.M. & Ritchie, M.J. (2012) ‘Roles of managers, frontline staff and local champions, in implementing quality improvement: stakeholders’ perspectives’, *Journal of Evaluation in Clinical Practice*, 18, pp.63–69.
- Kivimaki, M. & Elovainio, M. (1999) ‘A short version of the Team Climate Inventory: Development and Psychometric Properties’, *Journal of Occupational and Organizational Psychology*, 72(2), pp. 241-246.
- Korsen, N., Blount, A., Peek, C.J., Kathol, R., Narayanan, V., Teixeira, N., Freed, N., Moran, G., Newsom, A., Noda, J. & Miller, B.F. (n.d), *Collect and Use Data for Quality Improvement*. Available at: <https://integrationacademy.ahrq.gov/playbook/collect-and-use-data-quality-improvement>‘ (Accessed Sep 3, 2017).
- Kotter, J. P. (1995). ‘Leading change: why transformational efforts fail’, *Harvard Business Review*, March–April, pp. 59–67.
- Kriegel, R. J. (1991) *If it ain’t broke, break it!* Warner Books.
- Laozi (n.d.) *Wikipedia*. Available at:https://en.wiktionary.org/wiki/a_journey_of_a_thousand_miles_begins_with_a_single_step (Accessed Jun 4, 2016).

- Laureate Online Education (2010) ‘Week 9 The Doctoral Development Plan’
- Lee, L., Hillier, L.M., & Weston, W.W. (2014) ‘Ensuring the Success of Interprofessional Teams: Key Lessons Learned in Memory Clinics’, *Canadian Journal on Aging / La Revue canadienne du vieillissement*, 33(1), pp. 49 – 59.
- Lesser, E.L. (2000) *Knowledge and Social Capital*. Butterworth-Heinemann.
- Lester, H. & Campbell, S. (2010) ‘Developing Quality and Outcomes Framework (QOF) indicators and the concept of ‘QOFability’, *Quality in Primary Care*, 18(2), pp. 103–109.
- Lindenauer, P., Remus, D., Roman, S., Rothberg, M., Benjamin, E., Ma, A., & Bratzler, D. (2007) ‘Public reporting and pay for performance in hospital quality improvement’, *New England Journal of Medicine*, 356(5), pp. 486-496.
- Lugosi, P. (2006) ‘Between Overt and Covert Research: Concealment and Disclosure in an Ethnographic Study of Commercial Hospitality’, *Qualitative Inquiry*, 12(3), pp. 541-561.
- Lyngsnes, K. (2016) ‘A Reflexive eye on a collaborative action research project in school’, *The Qualitative Report*, 21(2), 196-211.
- Macpherson, A. & Jones, O. (2008) ‘Object-mediated learning and strategic renewal in a mature organization’, *Management Learning*, 39(2), pp. 177-201.
- Macpherson, A., Jones, O. & Oakes, H. (2006) *Mediating Artifacts, Boundary Objects and the Social Construction of Knowledge*. Citeseerx. Available at:
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.332.905> (Accessed Mar 30, 2016).
- Marcus, G.E. (1994) ‘On Ideologies of Reflexivity in Contemporary Efforts to Remake the Human Sciences’, *Poetics Today*, 15 (3), pp. 383-404.

Marsteller, J.A., Woodward, P., Underwood, W.S., Hsiao, C-J. & Barr, M.S. (2011) 'Design of a quality and performance improvement project for small primary care practices: reflections on the Center for Practice Innovation', *Quality in Primary Care*, 19, pp.49–57.

Martin, D. (2013) 'Scaling Up', *Danielle Martin at TEDxStouffville*. Jul 20. Available at: <https://www.youtube.com/watch?v=xa4PNZlmQSo> (Accessed Aug 7, 2017).

McDowall, A. & Saunders, M.N.K. (2010) 'UK managers' conceptions of employee training and development', *Journal of European Industrial Training*, 34(7), pp.609-630.

McGrail, K., Katz, A., Glazier, R. (2015) *Primary Health Care Performance Portraits*. Available at <http://www.transformationphc.ca/studies/study-1/> (Accessed Nov 25, 2015).

Mellahi, K., & Wilkinson, A. (2010). 'Managing and coping with organizational failure: Introduction to the special issue', *Group & Organization Management*, 35(5), pp. 531541.

Misner, I. (2011) *Four Steps to Building Social Capital*. Available at: <https://www.entrepreneur.com/article/219590> (Accessed August 17, 2017).

Mitchell, C (2015), *You Can't Make Me! - Effective Techniques for Managing Highly Resistant Clients*, Available at: <http://www.continuingcourses.net/active/courses/course036.php>, (Accessed May 21, 2017).

Morrison, E. W. & Milliken, F. J. (2000) 'Organizational silence: a barrier to change and development in a pluralistic world', *Academy of Management Review*, 25 (4), pp. 706–725.

Mulder, C., Glazier, R. & Sullivan, F. (2016) 'Ontario Data Support Starfield's Theory on Practice Quality and Cost', *Practice Based Research Network Annual Conference*.

Bethesda, USA, Jul 11-12. Available at:

<http://www.napcrg.org/Portals/51/Documents/PBRN%20Meeting/2016%20Meeting/Abstracts%202016%20OP%20FINAL.pdf?ver=2016-08-03-080525-080> (Accessed Nov 4, 2017).

Myrvold, S. (2011) '7 tips for continuous improvement', *American Printer*, 128 (7), pg. 9.

Nerenz, D.R. (2009) 'Ethical issues in using data from quality management programs', *Eur Spine J*, 18(Suppl 3), pp. S321–S330.

New Zealand. Internal Affairs (2015) 'Engagement strategy template', Available at:

<https://webtoolkit.govt.nz/guidance/online-engagement/engagement-strategyhttps://webtoolkit.govt.nz/guidance/online-engagement/engagement-strategy-template/-Content-strategytemplate/#Content-strategy> (Accessed May 21, 2017).

Nicolini, D., Mengis, J. & Swan, J. (2012) 'Understanding the Role of Objects in Cross-Disciplinary Collaboration', *Organization Science*, 23(3), pp. 612–629.

O'Brien, R. (2001) 'An Overview of the Methodological Approach of Action Research' in Richardson, R. (ed.), *Theory and Practice of Action Research*, João Pessoa, Brazil: Universidade Federal da Paraíba (English version), Available at <http://www.web.ca/~robrien/papers/arfinal.html> (Accessed Sep 12, 2017).

Ontario. Legislative Assembly of Ontario (2016) 'Patients First Act'. Toronto (Bill 41). Available at: http://www.ontla.on.ca/web/bills/bills_detail.do?locale=en&BillID=4215 (Accessed Jan 17, 2017).

Organisation for Economic Co-operation and Development (OECD), (2008) *Handbook on constructing composite indicators: Methodology and user guide*. Available at:

<http://www.oecd.org/els/soc/handbookonconstructingcompositeindicatorsmethodologyanduserguide.htm> (Accessed: Oct 13, 2014).

Oreg, S. & Sverdlik, N. (2011). ‘Ambivalence Toward Imposed Change: The Conflict Between Dispositional Resistance to Change and the Orientation Toward the Change Agent’, *Journal of Applied Psychology*, 96(2), pp. 337–349.

Orlikowski, W.J. (1996) ‘Improvising organizational transformation over time: A situated change perspective’, *Information Systems Research*, 7, pp. 63-92.

Owler, K., Morrison, R., & Plester, B. (2010) ‘Does fun work? The complexity of promoting fun at work’, *Journal of Management and Organization*, 16, pp.338-352.

Palmer, I. & Dunford, R. (2008) ‘Organizational change and the importance of embedded assumptions’, *British Journal of Management*, 19 (S1), pp. S20–S32.

Patients Canada (2015) *Performance Targets*. Available at:

<http://www.patientscanada.ca/index.cfm?id=51390> (Accessed Jul 7, 2017).

Patton, M.Q. (1999) ‘Enhancing the quality and credibility of qualitative analysis’, *Health Services Research*, 34(5), pp. 1189-1208.

Patton, M.Q. (2011) *Developmental Evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.

Pauchant, T.C. & Mitroff, I.I. (1988) Crisis prone versus crisis avoiding organizations: Is your company’s culture its own worst enemy in creating crisis? *Organization & Environment*, 2 (1), pp. 53–63.

Pava, M.L. (2008) ‘Why Corporations Should Not Abandon Social Responsibility’, *Journal of Business Ethics*, 83, pp. 805–812.

Pearson, C.M. & Clair, J.A. (1998) ‘Reframing crisis management’, *Academy of Management*

Review, 23 (1), pp. 59–76.

Pearson, S.D. & Raeke, L.H. (2000) ‘Patients’ trust in physicians: Many theories, few measures and little data’, *Journal of General Internal Medicine*, 15, pp. 509-513.

Perlman, K. (2011) *Change Fatigue: Taking Its Toll on Your Employees?* Available at:

<https://www.forbes.com/sites/johnkotter/2011/09/15/can-i-use-this-method-for-changehttps://www.forbes.com/sites/johnkotter/2011/09/15/can-i-use-this-method-for-change-in-my-organization/-533a5e8b1ce6in-my-organization/#533a5e8b1ce6>

(Accessed Oct 21, 2017).

Perneger, T.V. (2004) ‘Why we need ethical oversight of quality improvement projects’, *International Journal For Quality In Health Care*, 16(5), pp. 343-344.

Perry, C. and Zuber-Skerritt, O. (1994), “Doctorates by action research for senior practising managers”, *Management Learning*, 25 (2), pp. 341-64.

Piderit, S.K. (2000) ‘Rethinking resistance and recognizing ambivalence: A multi- dimensional view of attitudes toward an organizational change’, *Academy of Management Review*, 25 (4), pp. 783–794.

Platt, J. (1973) ‘Social traps’, *American Psychologist*, 28(8), pp. 641-651.

Platt, R., Grossmann, C. & Selker, H. (2013) ‘Evaluation as part of operations: Reconciling the common rule and continuous improvement’, *Hastings Center Report*, 43, pp. S37-9.

Prager, K. (2012) *Understanding behaviour change*. Available at:

<http://www.environment.scotland.gov.uk/media/16539/Understanding-Behaviourhttp://www.environment.scotland.gov.uk/media/16539/Understanding-Behaviour-Change.pdfChange.pdf> (Accessed Jun 3, 2016).

Premji, K.& Hogg, W. (2016) *Let's not confuse measurable with meaningful in primary care.*

Available at: http://healthydebate.ca/opinions/measure_performance_primary-care

(Accessed Dec 5, 2016).

Prochaska, J. O. & DiClemente, C. C. (1992) 'Stages of change in the modification of problem behaviors', *Progress in Behavior Modification*, 28, pp. 183-218.

Qaseem, A., Chou, R., Humphrey, L., & Shekelle, P. (2014) 'Inpatient glycemic control: best practice advice from the clinical guidelines committee of the American college of physicians', *American Journal of Medical Quality*, 29(2), pp. 95-98.

Raghunathan, T.E., (2004) 'What do we do with missing data? Some options for analysis of incomplete data', *Annual Review of Public Health*, 25, pp. 99–117.

Randall, J. (2004) *Managing change, changing managers*. Routledge

Reason, P. (2006) 'Choice and Quality in Action Research Practice', *Journal of Management Inquiry*, 15(2), pp. 187-203.

Reeves, D., Campbell S.M., Adams J., Shekelle P.G., Kontopantelis E. & Roland M.O. (2007) 'Combining multiple indicators of clinical quality: an evaluation of different analytic approaches', *Medical Care*, 45, pp. 489–496.

Revans, R.W. (1981) 'The nature of action learning', *Omega*, 9 (1), pp. 9-24.

Roberts, C. & Coghlan, D. (2011) 'Concentric collaboration: a model of leadership development for healthcare organizations', *Action Learning: Research and Practice*, 8(3) pp. 231-252.

Rolfe, G., Freshwater, D. & Jasper, M. (2001) *Critical reflection in nursing and the helping professions: a user's guide*. Basingstoke: Palgrave Macmillan.

Rosenstock, I.M. (1974) 'Historical origins of the health belief model', *Health Education Monographs*, 2, pp. 328-335.

Ryan, T. (n.d.) *When you reflect, are you also being reflexive?* Available at:

<http://oar.nipissingu.ca/PDFS/V812E.pdf> (Accessed Aug 14, 2017).

Sagarin, B.J., Ambler, J.K. & Lee, E.M. (2014) ‘An ethical approach to peeking at data’,

Perspectives on Psychological Science, 9(3), pp. 293–304.

Sandelowski, M. & Barroso, J. (2003) ‘Writing the Proposal for a Qualitative Research

Methodology Project’, *Qualitative Health Research*, 13(6), pp. 781-820.

Schein, E. H. (1999) ‘Kurt Lewin’s change theory in the field and in the classroom: notes toward

a model of managed learning’, *Reflections*, 1 (1), pp. 59-74.

Scholle, S.H., Roski, J., Adams, J.L., Dunn, D.L., Kerr, E.A., Dugan, D.P.& Jenson, R.E. (2008)

‘Benchmarking physician performance: reliability of individual and composite

measures’, *The American Journal of Managed Care*, 14 (12), pp. 833–838.

Schwarzer, R. (2008) ‘Modeling health behavior change: How to predict and modify the

adoption and maintenance of health behaviors’, *Applied Psychology: An International*

Review, 57(1), pp. 1–29.

Sezginalp, K (n.d.) *LGMA Membership Review: Strategies and Techniques to Attract and Retain*

Members, Available at:

http://www.lgma.ca/assets/Resources~and~Publications/Documents/Membership~Review~201/lgma_membership_review%20.pdf, (Accessed Sep 14, 2017).

Smith, D. & Elliott, D. (2007) ‘Exploring the barriers to learning from crisis: Organizational

learning and crisis’, *Management Learning*, 38(5), pp. 519-538.

Smith, K., & Berg, D. 1987. *Paradoxes of group life*. San Francisco: Jossey-Bass.

Smith, P.A.C. (1997) ‘Questioning action learning: More on minding our Ps and Qs’,

Management Decision, 35 (5), pp. 365-372.

- Southey, G. & Heydon, A. (2014) 'The Starfield Model - Measuring Comprehensive Primary Care for System Benefit', *Healthcare Management Forum*, 27 (2), pp. 60 – 64.
- Starfield, B. (1998) *Primary care: balancing health needs, services, and technology*, Oxford University Press.
- Starfield, B. (2009) 'Toward international primary care reform', *Canadian Medical Association Journal*, 180(11), pp. 1091-2.
- Staw, B.M. & Ross, J. (1987) 'Knowing when to pull the plug', *Harvard Business Review*, 65, pp. 68-74.
- Stempniak, M. (2013) 'Quality control: Doctors lobby for adaptable quality measures', *Hospitals & Health Networks*, 87 (8), p. 14.
- Stetzer, E (2014) *Strategic Evangelism: The Power of an Invitation*. Available at:
<http://www.christianitytoday.com/edstetzer/2014/july/power-of-invitation-our-god>
<http://www.christianitytoday.com/edstetzer/2014/july/power-of-invitation-our-god-pursues-lost-and-so-should-we.html>, (Accessed Sep 14, 2017).
- Taylor, C., Bury, M., Campling, N., Carter, S., Garfield, S., Newbould, J. & Rennie, T. (2006), *A Review of the use of the Health Belief Model (HBM), the Theory of Reasoned Action (TRA), the Theory of Planned Behaviour (TPB) and the Trans-Theoretical Model (TTM) to study and predict health related behaviour change*. Available at:
<https://www.nice.org.uk/guidance/PH6/documents/behaviour-change-taylor-et-al>
<https://www.nice.org.uk/guidance/PH6/documents/behaviour-change-taylor-et-al-models-review2models-review2> (Accessed Feb 2, 2017).
- Taynor, C.B. (2007) *Six Sigma Software Development*. 3rd edition. Auerbach Publications.

Tews, M.J., Michel, J.W. & Allen, D.G. (2014) 'Fun and friends: The impact of workplace fun and constituent attachment on turnover in a hospitality context', *Human Relations*, 67(8), pp. 923-946.

The Conference Board of Canada (2014) *Final Report: An external evaluation of the Family Health Team (FHT) initiative*. Available at:

<http://www.conferenceboard.ca/ehttp://www.conferenceboard.ca/e-library/abstract.aspx?did=6711&AspxAutoDetectCookieSupport=1library/abstract.aspx?did=6711&AspxAutoDetectCookieSupport=1> (Accessed Feb 18, 2015).

Thomas, R., Sargent, L.D. & Hardy, C. (2011) 'Managing Organizational Change: Negotiating Meaning and Power-Resistance Relations', *Organization Science*, 22 (1), pp. 22–41.

Tschirhart, M. & Gazley, B. (2014) 'Advancing Scholarship on Membership Associations: New Research and Next Steps', *Nonprofit and Voluntary Sector Quarterly*, 43(2S), pp. 3S–17S.

Tsoukas, H. & Chia, R. (2002) 'On organizational becoming: rethinking organizational change', *Organization Science*, 13 (5), pp. 567-582.

Turnbull, S. (2001) 'Corporate Ideology – Meanings and Contradictions for Middle Managers', *British Journal of Management*, 12(3), pp. 231-242.

Turner, B.A. (1976) 'The organizational and interorganizational development of disasters', *Administrative Science Quarterly*, 21 (3), pp. 378–397.

Van de Ven, A.H. & Poole, M.S. (1995) 'Explaining development and change in organizations', *Academy of Management Review*, 20 (3), pp. 510–540.

Van Winkel, C. (2016) 'Using developmental evaluation methods with communities of practice', *The Learning Organization*, 23(2/3), pp.141-155.

- Vince, R. & Broussine, M. (1996) ‘Paradox, defense and attachment: accessing and working with emotions and relations underlying organizational change’, *Organization Studies*, 17 (1), pp. 1–21.
- Wagner, W.G., Pfeffer, J. & O'Reilly III, C.A. (1984) ‘Organizational Demography and Turnover in Top-Management Groups’, *Administrative Science Quarterly*, 29(1), pp. 74–92.
- Weick, K.E. (1985) ‘Cosmos vs. Chaos: Sense and Nonsense In Electronic Contexts’, *Organizational Dynamics*, 14(2), pp. 51-64.
- Weick, K.E. (1988) ‘Enacted sense-making in crisis situations’, *Journal of Management Studies*, 25 (4), pp. 305–317.
- Weick, K.E. & Quinn, R.E. (1999) ‘Organizational change and development’, *Annual Review of Psychology*, 50, pp. 361–386.
- Westerman, M.A. (2006) ‘What counts as “good” quantitative research and what can we say about when to use quantitative and/or qualitative methods?’, *New Ideas in Psychology* 24, pp. 263–274.
- Whicher, D., Kass, N., Saghai, Y., Faden, R., Tunis, S. & Pronovost, P. (2015) ‘The views of quality improvement professionals and comparative effectiveness researchers on ethics, IRBs, and oversight’, *Journal of Empirical Research on Human Research Ethics*, 10 (2), pp. 132-144.
- Wiest, D.A. (2006) ‘The Challenges of a Change Agent’, *Top Emerg Med*, 28(2), pp. 125–128.
- Wilkinson, S. (1988) ‘The Role of Reflexivity in Feminist Psychology’, *Women's Studies International Forum*, 11, pp. 493–502.

- Wolfson, D., Bernabeo, E., Leas, B., Sofaer, S., Pawlson, G. & Pillittere, D. (2009) 'Quality improvement in small office settings: an examination of successful practices', *BMC Family Practice*, 10 (14), pp. 1-12.
- Yardley, S.J., Watts, K.M., Pearson, J. & Richardson, J.C. (2014) 'Ethical issues in the reuse of qualitative data: Perspectives from literature, practice, and participants', *Qualitative Health Research*, 24(1), pp. 102–113.
- Yeo, R.K. (2013) 'Building commitment through reflective practice: an inside-out approach', *Industrial and Commercial Training*, 45 (1), pp. 3 – 11.
- Zaslavsky, A.M., Shaul, J.A., Zaborski, L.B., Cioffi, M.J. & Cleary, P.D. (2002) 'Combining health plan performance indicators into simpler composite measures', *Health Care Finance Review*, 23 pp. 101-115.