Completing the Feedback Loop: An Assessment of the Implementation of an Early Stage Audit and Feedback Intervention in Primary Care

by

Daniel Jonathan Wagner

A thesis submitted in conformity with the requirements for the degree of Master of Science in Health Services Research

Institute of Health Policy, Management and Evaluation
University of Toronto

© Copyright by Daniel J. Wagner 2016
Completing the Feedback Loop: An Evaluation of the Implementation of an Early Stage Audit and Feedback Intervention in Primary Care

Daniel Jonathan Wagner
Master of Science in Health Services Research
Institute of Health Policy, Management and Evaluation
University of Toronto
2016

Abstract

Audit and Feedback (A&F) is a common quality improvement strategy characterized as an feedback loop enabling an iterative, self-regulating progress. The objective of this thesis was to evaluate the implementation of a specific A&F program in team-based primary care practices. Eighteen semi-structured interviews were conducted to explore its implementation and factors influencing participation. Interview transcripts were analyzed deductively with the support of the Consolidated Framework for Implementation Research. Findings revealed that implementation was not representative of a complete feedback loop. Participation was primarily associated with a perception that the program would develop into a best-in-class A&F initiative for primary care. The relationship between physicians and organizational administrators was identified as a key factor which may support or deter implementation. The study also identified factors related to the design and delivery of A&F that may explain why and how an A&F intervention in primary care can achieve its objectives.
Acknowledgements

The completion of this thesis project would not have been possible without the many partnerships I was fortunate enough to establish in the last eight-months. I am especially grateful to Carol Mulder and her colleagues at the Association of Family Health Teams of Ontario (AFHTO) for their trust and support in the completion of this thesis. It is my hope that this project can be used to support AFHTO on its mission to create an audit and feedback tool which is meaningful and supportive of the work done by Family Health Teams.

I would also like to thank Dr. Holly Witteman at Université Laval and Dr. Laura Desveaux at Women’s College Hospital who supported my journey in a new area of research I had never imagined I would pursue. Having knowledgeable people available to review my methodology or to provide general feedback was invaluable in this process.

I am grateful to the support and guidance I received from my committee members, Dr. Jan Barnsley and Dr. Janet Durbin. Having seen other students go through their thesis projects, I know that the amount of time committed by them to my thesis is not the norm. This encouraged me to continue to push myself to deliver on my commitments and create a thesis of a high quality. I must also take a moment to thank Dr. Durbin, who was instrumental in my successful transition to the University of Toronto. While our original partnership did not work out as planned, I am appreciative that we were able to extend our partnership into this project.

I must also extend my gratitude to Dr. Charles Popper of Harvard University whose mentorship has been invaluable over the last three years. I consider myself very fortunate to have so many established leaders, such as Dr. Popper, supporting my professional development.
The success I have achieved in my graduate studies would not have been possible without the mentorship I received from Dr. John McLennan at the University of Calgary. There is perhaps no better sign of a strong mentor than what their pupils accomplish next. My success in my graduate studies was enabled from the preparation I received from Dr. McLennan. I am extremely grateful that our partnership has continued throughout my graduate work as we successfully published two peer-reviewed articles.

I’d also like to thank Dr. Noah Ivers for taking a chance on me and supporting me in developing and executing this project in eight-months. Having worked with other investigators at other institutions, I am aware that my relationship with Dr. Ivers is unique. For this thesis, I have been supervised by a mentor whose enthusiasm for Health Services Research and Audit and Feedback is infectious. This drove me to not only be ambitious with what I hoped to achieve with this project, but gave me the confidence to pursue a new research methodology which I never considered attempting previously. I can think of no better learning environment.

Through their limitless encouragement and support, my family has played an integral role which has enabled me to continue to push the boundaries of my success. I am fortunate to have such great role models from whom I can continue to look for guidance.

Finally, I’d like to extend my deepest gratitude to Calgary Academy – a school which changed my life. In one short year, I was taught how to overcome my challenges and leverage my strengths to be successful. I can think of no larger contributor to my current and continued success. However, I am humbled knowing my case was unique. I continue to work hard in order to set an example for those who face similar challenges as I did.
# Table of Contents

Abstract .................................................................................................................. ii
Acknowledgements .................................................................................................. iii
List of Tables .......................................................................................................... vii
List of Plates .......................................................................................................... viii
List of Figures ......................................................................................................... ix
List of Appendices ................................................................................................. x

Background ........................................................................................................... 1

Primary Care .......................................................................................................... 1
Measuring the Quality of Primary Care ................................................................... 5
Audit and Feedback ............................................................................................... 7
Types of Audits ....................................................................................................... 8
Conceptualizing the Intervention .......................................................................... 9
Effectiveness of Audit and Feedback .................................................................... 10
Audit and Feedback in Primary Care .................................................................... 11

Gaps in the Literature ............................................................................................ 20

Objective ............................................................................................................... 22

Methods .................................................................................................................. 23

Basic Research Strategy ....................................................................................... 23

Intervention ............................................................................................................ 25

Practice Characteristics ....................................................................................... 30
Safe ....................................................................................................................... 31
Effective ................................................................................................................ 31
Patient-Centred .................................................................................................... 31
Timely .................................................................................................................... 32
Efficient ................................................................................................................ 32
Consolidated Quality Measure .......................................................................... 32

Theoretical Framework ......................................................................................... 35

Measures ............................................................................................................... 38
D2D Website Analytics ......................................................................................... 38
D2D 3.0 Data ........................................................................................................ 39
Qualitative Data .................................................................................................... 39

Analysis ................................................................................................................. 42
Question 1 .............................................................................................................. 44
Question 2 .............................................................................................................. 45
List of Tables

Table 1. Identified Features of Audit and Feedback Interventions in Primary Care .................. 15
Table 2. Core Measures of the D2D Audit and Feedback Report .............................................. 34
Table 3. Domains of the Consolidated Framework for Implementation Research ...................... 37
Table 4. Summary of Interview Format and Participants ........................................................... 48
Table 5. Summary of Core Measures for D2D 3.0 ..................................................................... 50
Table 6. Summary of FHT Implementation of D2D by CFIR Domain .......................................... 57
Table 7. Summary of Facilitators in Deciding to Participate in D2D by CFIR Domain ................. 70
Table 8. Summary of Barriers in Implementation of D2D by CFIR Domain ............................... 92
List of Plates

Box 1. Data Submission ........................................................................................................ 52
Box 2. Receipt of Feedback .................................................................................................. 55
Box 3. Act on Feedback ........................................................................................................ 56
Box 4. Peer Comparison and Benchmarking ...................................................................... 59
Box 5. Perceived Recency ..................................................................................................... 59
Box 6. Multi-Source Audit .................................................................................................... 60
Box 7. Promise of Future Potential ........................................................................................ 61
Box 8. Informed by Participant Feedback ............................................................................. 62
Box 9. Evidence Base for D2D ............................................................................................ 63
Box 10. Advise Policy on Direction of QI in Primary Care .................................................. 64
Box 11. Advise Policy on the Value of Family Health Teams ............................................. 65
Box 12. Peer Influence .......................................................................................................... 66
Box 13. Role as AFHTO Member .......................................................................................... 66
Box 14. Availability of Implementation Leader .................................................................... 67
Box 15. Leadership Role at AFHTO ..................................................................................... 68
Box 16. Development of QI Capacity ................................................................................... 69
Box 17. Frequency of Audit Cycles ..................................................................................... 72
Box 18. Timing of Measures ................................................................................................. 73
Box 19. Sensitivity ................................................................................................................ 75
Box 20. Duplication .............................................................................................................. 76
Box 21. Definitions ................................................................................................................ 77
Box 22. Peer Comparison ..................................................................................................... 78
Box 23. Website Design ........................................................................................................ 80
Box 24. QIDSS Dependence ............................................................................................... 82
Box 25. Relative Priority ....................................................................................................... 85
Box 26. Time .......................................................................................................................... 87
Box 27. Human Resources ................................................................................................... 89
Box 28. FHT-FHO Relationship .......................................................................................... 91
List of Figures

Figure 1. Basic Research Strategy ............................................................... 24
Figure 2. Timeline of Data and Release of the Data-2-Decisions Audit and Feedback Report ... 27
Figure 3. Screenshot of Web-Based Interactive Feedback Report for D2D. ........................ 29
Figure 4. Summary of Participant Recruitment for Interviews ............................... 47
Figure 5. Summary of Website Activity Between February 1 to March 14, 2016 ............... 53
List of Appendices

Appendix 1. Ethics Documentation .................................................................................. 113
Appendix 2. D2D Data Submission Form ....................................................................... 116
Appendix 3. Interview Recruitment Letters .................................................................... 117
Appendix 4. Interview Registration Form ........................................................................ 119
Appendix 5. Qualitative Interview Guide ....................................................................... 120
Appendix 6. Qualitative Codebook .............................................................................. 126
Background

Primary Care

Efforts to improve primary health care can be traced back to the declaration of Alma-Ata in 1978, which was the first occasion that the importance of primary health care was emphasised by the international community (1). The declaration outlined five principles of primary health care: equity, participation, inter-sectoral action, appropriate technology, and a central role played by the health system (1). One element of primary health care is the delivery of Primary Care (PC) services.

There is much confusion regarding the definition and scope of primary care. This may be attributable to the absence of a single universally applicable definition to describe the role and function of Primary Care services. Many attempts have been made by a variety of institutions and people to define primary care. For example, the World Health Organization (WHO) defines PC as “a key process in the health system which includes first-contact, accessible, comprehensive and coordinated care” (2). Meanwhile, the Canadian Medical Association (CMA) defines PC as “first-contact assessment of a patient and the provision of continuing care for a wide range of health concerns” (3). The definition proposed by the Institute of Medicine (IOM) is slightly different, in which PC is defined as “the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health needs, developing a sustained partnership with patients and practicing in the context of family and community” (4).

In addition to the widely referred definitions proposed by the WHO, CMA and IOM, much attention has been given to Dr. Barbara Starfield’s conceptualization of primary care. In
her book, Dr. Starfield described PC as “the level of a health service system that provides entry into the system for all new needs and problems, provides person-focused (not disease oriented) care over time, provides care for all but very uncommon or unusual conditions, and co-ordinates or integrates care provided elsewhere by others” (5). This conceptualization of Primary Care seems to capture many of the themes raised in those by the WHO, CMA, and IOM while also framing the definition to be of utility for those interested in the study of health services research in PC, regardless of the structure of Primary Care services in each jurisdiction.

Over the last thirty years, evidence has been mounting which demonstrate the benefits of a health care system which is oriented towards PC rather than specialist services in terms of outcomes and health systems costs (6,7). Many studies have observed a relationship between the strength of a country’s primary care system and improved primary care outcomes. Specifically, countries with greater orientation towards primary care tend to have lower all-cause mortality, all-cause premature mortality, and even cause-specific premature mortality for some cardiovascular and respiratory conditions (8). It is argued that improved access to primary care physicians could yield reduced hospitalizations, decreased utilization of specialist and emergency resources, and a reduced likelihood of being subjected to inappropriate health interventions (6).

In Canada, access to primary care is strong with 85.1% of the population having regular access to primary care (9). However, Canada has a federalist health care system which provides federal transfer payments to each province in return for the provision of health care which is publicly administered, comprehensive, universal, portable, and accessible (10). As a result, the Canadian health system is really a composition of 10 provincial and 3 territorial health systems. In examining access to primary care by province, considerable variation in access to primary
care can be observed. For example, access is greatest in New Brunswick (93.9%) and lowest in Quebec (74.8%) while Ontario, Canada’s most populous province, is near the top of the list with 92.5% of its population with a regular primary care physician (9).

The majority of provinces in Canada have structured primary care services around a fee-for-service (FFS) model, where-in primary care physicians may choose to practice in solo or group settings. In Ontario, this model began to display its weaknesses as the costs of health care continued to rise and fewer medical graduates were practicing family medicine. Economic incentives directed PC providers to maximize the volume of patient visits, leading to increased utilization of health care services (11). However, FFS by its nature did not incentivize the features of PC which had been associated with more cost-efficient health systems (i.e., comprehensiveness, continuity, coordinated, accessible) (11).

Efforts by the Government of Ontario to initiate inter-professional, team-based primary care can be traced back to the introduction of Community Health Centres (CHC) in 1979 (11,12). This new primary care practice model brought together multidisciplinary care teams, and new forms of physician remuneration to primary care (11,12). It would take another twenty years before primary care reforms would be introduced more widely. Family Health Networks (FHN) and Family Health Groups (FHG), introduced in 2001 and 2003, expanded the use of new payment models and practice structures (12). These primary care models also paved the way for the introduction of the Family Health Team (FHT) in 2004 (12). FHTs are intended to support a multidisciplinary team of health service providers in the provision of high quality, patient-centred care to rostered Ontarians. Today, there are 294 FHTs operating in Ontario, serving more than 3.5 million people (13).
A Family Health Team is a combination of two legal entities, a Family Health Organization (FHO) and a FHT. The FHO represents the group of physicians who agree to work with the FHT. However, it is the FHT which is responsible for the expansion of the typical scope of primary health care services. This is accomplished through the employment of a variety of health professionals, also referred to as interdisciplinary health professionals (IHP). Examples of IHPs employed by a FHT include Nurse Practitioners, Registered Nurses, Pharmacists, Social Workers, Psychologists, Health Educators, and Occupational Therapists (12,14). All providers employed by the FHT are remunerated by the Ministry of Health and Long-Term Care (MoHLTC) and the number of each provider type is typically determined by the size of each FHT’s respective patient roster. Meanwhile, remuneration of physicians practicing at the FHT is determined by a blended funding model which emphasizes capitation, with a relatively small amount from fee-for-service and bonuses. This remuneration strategy is meant to incentivize appropriate management of chronically ill patients as well as population oriented preventive care (12,14).

The FHT serves as the site of most patient care for enrolled patients, commonly referred to as a patient’s “medical home”. The FHT model was developed at a similar time as literature began to emerge elsewhere regarding the development of medical homes and are thought to meet similar standards and requirements (14). To improve access, each practice is required to provide 3-hour evening and weekend clinics for walk-in and scheduled patients as well as an after-hours on call system (12,14). In addition, FHTs are required to submit “a formal, documented set of quality commitments aligned with system and provincial priorities that a health care organization makes to…improve quality through focused targets and actions” on an annual basis (15). Referred to as a Quality Improvement Plan (QIP), these are submitted to Health Quality Ontario
(HQO), the advisor to the province on health care quality (16). HQO was created in 2010 as one element of the Excellent Care for All Act, which established the mandate for the organization (17). Responsibilities include, i) to monitor and report on the performance of the health system; ii) to provide guidance on important quality issues, iii) to assess evidence to determine what constitutes optimal care; iv) to engage patients and give them a voice in shaping the health system; and v) to promote continuous quality improvement aimed at substantial and sustainable positive change in health care (18).

**Measuring the Quality of Primary Care**

In 2001, the Institute of Medicine outlined six domains of quality which a health system, or a component of a health system, should work to achieve. First, health care should be *safe*. This means that efforts should be made to avoid injuries to patients from the care that is intended to help them (19). Second, health care should be *effective*. Services, which provide evidence based treatments, should be delivered to all who could benefit but not to those unlikely to benefit (19). Third, health services should strive to be *patient-centred*. The provision of care should be respectful of and responsive to individual patient preferences, needs, and values and ensure that patient values guide clinical decisions (19). The fourth domain of quality is that health care should be *timely*. Efforts should be made to reduce waits and sometimes harmful delays for both those who receive care (19). Fifth, health care should be *efficient*. This quality domain refers to avoiding waste, including waste of equipment, supplies, ideas, and energy (19). The sixth and final domain of quality is that health care should be *equitable*. This refers to providing care that does not vary in quality because of personal characteristics such as gender, ethnicity, geography, or socioeconomic status (19).
While the six domains of quality provide an important evaluative framework, measuring the achievement of each goal in primary care is a complicated task. This is due to the complexity of the primary care practice setting. Unlike specialist care, primary care services are provided to respond to a wide range of issues crossing the boundaries of many specialties. For example, the measurement of the effectiveness of PC services faces many barriers when using published standards as a comparator. This may be attributable, in part, to the use of samples which exclude patients with comorbidities in the calculation of such standards. However, primary care providers typically see patients with more complex profiles, meaning they may take a more holistic approach to their care. This inherently makes each individual issue more complex to manage. An additional factor complicating the measurement of quality in primary care is the introduction of newer, team-based practice models – like the FHT in Ontario. It could be argued that the teamwork component may contribute to or undermine the quality of primary care. For example, a medical home or primary care team implies a relationship between the patient and a group of care providers, who are accountable to the team. This is in contrast to non-team models where the patient has a relationship with an individual primary care provider.

Heath et al., argue that effective primary care depends on the integration of both vertical and horizontal care. The former concerns the management of specific diseases from primary to tertiary care. Meanwhile, the latter emphasizes the integration of care and the needs of individuals and the design of systems of care that focus on the broad needs of the community and population (20). While the distinction between vertical and horizontal care is important, assessment of care in both domains remains a difficult and controversial task. The point of distinguishing between these levels of care is meant to demonstrate that quality of care needs to be assessed and calibrated at many different levels. These may include (but are not limited to),
the implementation of evidence-based best practice guidelines; the needs, values and priorities of patients; population level indicators; and indicators monitoring the status of the health system like affordability and equity (20).

To prevent undermining the achievement of quality between domains, a balance must be struck in performance measurement in primary care to ensure that too much attention is not given to a set of measures at one level (20). To reach such a balance, initiatives commonly referred to as *quality improvement* can provide focus to the measurement or improvement of the quality of primary care. Quality improvement may be defined as “the combined and unceasing efforts of everyone – healthcare professionals, patients and their families, researchers, payers, planners and educators – to make the changes that will lead to better patient outcomes (health), better system performance (care) and better professional development” (21). A wide range of strategies have been employed as part of quality improvement initiatives over time. Examples of interventions may include, adherence to clinical practice guidelines, peer review, recording consultations and patient interviews. One additional quality improvement tool which has garnered considerable attention and focus over time has been medical audit with feedback (22,23).

**Audit and Feedback**

The intention of an audit and feedback (A&F) intervention is to prompt changes by healthcare professionals and/or organizations to address gaps between ideal and expected care. A&F is most commonly defined as delivering to health providers “a summary of the clinical performance of healthcare provider(s) over a specified period of time” (24). The National Institute for Clinical Excellence (NICE) defines A&F as “a quality improvement process that seeks to improve patient care through systematic review of care against explicit criteria and the
implementation of change” (25). While the two definitions share similar features, here forward A&F will refer to the former definition rather than the latter as it is most used in the literature.

One problem commonly identified with the A&F literature is that there are many phrases which either have been or could be used to describe an A&F intervention. Examples include “medical record review”, “quality care review”, “medical process evaluation”, and “quality assurance”. As a result, a definition of A&F alone may be insufficient to explore the current knowledge in the field without clarifying the different methodologies which can be used in an A&F intervention. Therefore, it is important to distinguish between the features of audit strategies commonly employed.

**Types of Audits**

While a variety of different audits are utilized to assess medical practice, they can be classified into one of three types. First, *external audits* are those conducted by individuals who are external to the setting which is the subject of the audit (26,27). Second, *peer audits* are conducted by small groups of peers, often working in the same practice, who agree on setting standards and realistic objectives in assessing the quality of care they provide (26,27). Third, *self-audits* are those which are conducted by the provider themselves in order to improve the care they provide (26,27). The result of each type of audit is to provide feedback in order to improve some part of the practice.

Given that audits can examine different elements of the quality continuum, it can be useful to assign categories to the elements of practice which the audit seeks to assess. First, *audit of structure* refers to the quality and types of services available at the practice. This may include administrative processes, collaboration within the practice and specialists and practice management, and provider to patient ratios (26,27). Second, *audit of process* refers to what is
actually done to the patient. Processes can be divided into clinical performance (examination, referrals, prescriptions), interpersonal performance, and managerial performance (27). The most common methodology employed is a chart audit, whereby a limited number of quality indicators are selected which are deemed critical to the process of care (26). Third, audit of outcome is an assessment of the eventual results of the intervention.

While audits of outcome are often considered to be the most relevant indicators of the quality of patient care, they are often paired with audits of process (26,27). The rationale for this type of evaluation is to ensure that the results of the audit process provide a valid reflection of the quality of care (26,27). For example, it is well understood that factors beyond medical care will influence patient outcome such as income, the patient’s behaviour in maintaining their health, or the environment in which they live (26,27). Absence of the consideration of such factors on the outcome of care may lead to an imperfect understanding of the quality of care provided when the results of the audit are disseminated to providers during the feedback portion of the audit cycle.

**Conceptualizing the Intervention**

As a formal quality improvement strategy, Audit and Feedback may be conceptualized by using the methodologies of medical audit, described in the previous section, to summarize performance and eventually improve it. A&F is characterised by a feedback loop, which represents an iterative self-regulating process (28). First, some measure (or series of measures) are monitored over time to capture “current” performance. These results are then compared with some type of performance target. Examples include published benchmarks in clinical practice guidelines, or the average scores achieved by the top performers in a practice’s geographic region. Second, information on the comparative level of performance is fed back by an external
source to the actor. Specifically, feedback may be provided using the postal service, electronic mail, in-person review, or through a computer based service (such as an internet based web portal or an internal practice dashboard). Third, action plans detailing specific behaviours required to improve performance to achieve the performance target are then articulated. In some circumstances, the subject receiving the feedback may also be aided in the development of such initiatives (28). Once the action plan has been developed, this cycle is completed again. It should be noted that the type of audit utilized may influence the audience for the feedback data as well as the actors. For example, process audits may be of utility to both administrators and clinicians who may need to work together to re-calibrate clinical pathways to improve clinical efficiency, if necessary. Meanwhile, outcome audits may be of utility to clinicians who may wish to change a specific practice behaviour to ensure they continue to meet the objective of treatments they’re providing.

**Effectiveness of Audit and Feedback**

The effectiveness of A&F was most recently evaluated in the third update of a Cochrane review, which included 140 randomized trials of audit and feedback interventions conducted across many clinical conditions and settings around the world. With respect to process measures, audit and feedback yielded a median improvement of 4.3% (interquartile range = 0.5% to 16%). Meanwhile, the effectiveness of audit and feedback influencing patient outcome measures was less clear. Results for studies with dichotomous outcomes revealed a median of 0.4% (interquartile range -1.3% to 1.6%) while studies with continuous outcome measures demonstrated a median improvement of 17% (IQR 1.5 to 16%). The discrepancies in the effectiveness of A&F on patient outcome measures may be a function of the sample size in the
Cochrane review given that 12 comparisons from six studies were completed for the former and eight comparisons from five studies were completed to compute the latter measure.

In addition to summarizing the effectiveness of A&F using standard approaches, the new update of the Cochrane review also introduced much needed theoretical appraisals of the different components of A&F interventions. This is of value as it is argued that for behavioural interventions such as audit and feedback, theoretical appraisal offers a systematic framework for categorizing and evaluating intervention content. Specifically, the application of theory to evidence synthesis enabled the specification of intervention techniques and mechanisms by which any effects are achieved (24).

In using this approach within the most recent update of the Cochrane review on Audit and Feedback Interventions, multivariate meta-regression analysis was used to identify five factors which may improve the success of an A&F intervention. First, success may be more likely when baseline performance is considered “low”. Second, the intervention may be effective when the source of feedback is a supervisor or colleague. Third, increased frequency of feedback may yield improvement compared to one feedback session. Fourth, A&F may be more successful if delivered in both verbal and written formats. Lastly, audit and feedback may be successful if it includes both explicit targets and an action plan (24).

Audit and Feedback in Primary Care

For decades, Audit and Feedback has been used as a quality improvement intervention across the entire spectrum of health services. Examples include Emergency Departments, inpatient hospital wards, intensive care units, and outpatient specialty services. It has also been implemented in primary care in various forms. A literature review of articles focusing on the subject of Audit and Feedback, conducted for this thesis, identified ninety-one articles across the
spectrum of health services. Of those identified, twenty-two explored A&F in primary care in the form of literature reviews, commentaries, or primary investigations using qualitative or quantitative methodologies.

Two qualitative studies were identified which investigated features of high and low performing practices as well as how A&F is used in primary care. First, a 2006 study by Hysong et al. investigated how high and low performing practices differ in the way in which they use results from an A&F cycle (29). Practices were identified based on their adherence to clinical practice guideline recommendations for six common chronic conditions and the three highest and three lowest performing practices were selected for interviews. Interviews were conducted with a variety of members of the practice including leadership, middle management and support staff, and clinic staff (29).

Results from the interviews identified four characteristics of feedback delivery. These include timely, individualized, punitive, and customizable. Timely feedback (defined as monthly or more frequent) was consistently observed for high performing practices compared to low performers (29). Individualized data was a typical characteristic observed among high performers. Meanwhile, low performers were more likely to use aggregated data – which is more difficult for a provider to act on. Findings from interviews revealed that high performers would explicitly state that results of feedback would not result in punishment (29). Lastly, feedback which was customizable in its presentation to the provider was a typical characteristic of high performing facilities (29).

Second, a 2014 study by Ivers et al. sought to assess the perceived utility and to identify barriers to family physicians’ use of A&F (30). Each of the participating physicians had received one feedback report detailing their performance on nine evidence-based quality indicators using
Data abstracted from their respective electronic medical records (30). Participants were selected using stratified purposeful sampling, selecting participants with features reported as relevant in previous studies with the aim of seeking informational representativeness. A total of twelve interviews were conducted with physicians practicing in the Family Health Team model (30).

Results from the interviews revealed that no physician found the feedback to be useful. Respondents were concerned about the validity of the data used to generate the reports and the ability to leverage EMRs to support quality improvement efforts (30). Other concerns included managing priorities between standardized quality targets and patient centered care, and appropriate use of physician time to engage properly in quality improvement activities. Physicians indicated mixed preferences for interventions designed to support quality improvement initiatives. A final barrier identified by participants was the absence of infrastructure to support QI in each practice (30).

In addition to the qualitative studies discussed above, 13 quantitative studies were identified which evaluated specific A&F interventions. As demonstrated in Table 1, these studies reflect a wide range of inconsistencies in the ways in which A&F has been deployed in Primary Care. This observation should not come as a surprise given the complexity of monitoring quality in Primary Care and the broad scope of health issues which may be treated. However, several trends were identified in the way in which A&F has been deployed in PC. For example, the majority of interventions could be categorized as process audits rather than outcome audits. Feedback was most commonly delivered to the physician alone and least frequently targeted towards all clinic staff. Quarterly feedback cycles were most common and annual feedback cycles were observed to be the least common among the studies reviewed. Data abstraction was most frequently completed via manual chart review (either paper chart or an electronic medical
record). Lastly, the scope of the audit itself varied widely from very specific outcome measures to process measures (i.e. cancer screening).
Table 1. Identified Features of Audit and Feedback Interventions in Primary Care

<table>
<thead>
<tr>
<th>Audit Type</th>
<th>% (n)</th>
<th>Audit Source</th>
<th>% (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Audit Only</td>
<td>69.2 (9)</td>
<td>Electronic Medical Record</td>
<td>69.2 (9)</td>
</tr>
<tr>
<td>Outcome Audit Only</td>
<td>15.4 (2)</td>
<td>Paper Medical Record</td>
<td>38.5 (5)</td>
</tr>
<tr>
<td>Both</td>
<td>15.4 (2)</td>
<td>Patients</td>
<td>15.4 (2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feedback Target Audience</th>
<th>% (n)</th>
<th>Frequency of A&amp;F Cycles²</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD</td>
<td>100 (13)</td>
<td>One time</td>
</tr>
<tr>
<td>Nurse</td>
<td>30.8 (4)</td>
<td>Quarterly</td>
</tr>
<tr>
<td>Admin</td>
<td>30.8 (4)</td>
<td>Bi-Annually</td>
</tr>
<tr>
<td>Other Clinic Staff</td>
<td>23.1 (3)</td>
<td>Annually</td>
</tr>
<tr>
<td>All Staff</td>
<td>23.1 (3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Abstraction³</th>
<th>% (n)</th>
<th>Data Abstraction: Manual Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated EMR Process</td>
<td>30.8 (4)</td>
<td>All Eligible Charts</td>
</tr>
<tr>
<td>Manual Review</td>
<td>61.5 (8)</td>
<td>Random Sample</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A&amp;F Scope</th>
<th>% (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height/Weight/Vision/Hearing/BP</td>
<td>23.1 (3)</td>
</tr>
<tr>
<td>Cardiovascular Disease</td>
<td>23.1 (3)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>38.5 (5)</td>
</tr>
<tr>
<td>Cancer Screening</td>
<td>23.1 (3)</td>
</tr>
<tr>
<td>Immunizations</td>
<td>38.5 (5)</td>
</tr>
<tr>
<td>Access to Primary Care</td>
<td>7.7 (1)</td>
</tr>
</tbody>
</table>

Notes:
1. Cited Articles: (31–43)
2. Frequency of A&F Cycles was not specified in 1 article.
3. Data abstraction procedure not specified in 1 article.
While the above demonstrates the diversity of approaches utilized to implement Audit and Feedback in Primary Care, a closer examination of the studies and their implications is necessary to understand the utility of implementing A&F in primary care. For example, A&F may be used to monitor and improve access in primary care. A 2010 study conducted an audit of structure to measure and improve the accessibility and availability of general practice. Utilizing a pre-post design, questionnaires measuring the outcomes of interest were completed by patients and providers at baseline and follow-up. In addition, a “mystery” patient was hired to contact and seek care in the participating practices at random times throughout the study. Feedback from baseline results were summarized and delivered to each practice for the purposes of developing an action plan to facilitate improvement. While the study lacked a control group, and was conducted over a short time frame, results indicated a 10% improvement in phone accessibility in the participating practices. Further, 80% of the action plans drafted had been completed and implemented after a period of only 5-months (37).

Audit and feedback may also be used for quality improvement of more unique cases, which are not typically seen in primary care. In this case, a study by Raasch et al. implemented A&F to improve care of suspicious skin lesions in primary care. The authors conducted a prospective randomized trial in two distinct practice settings located within a remote region of Australia where these cases are more commonly seen in primary care. While results suggested that providers’ confidence in diagnosis and management did improve, the two groups were too different to assess the outcomes with certainty (41).

A&F interventions have also been used to improve medication prescribing patterns in primary care. For example, a 2013 study investigated the extent of medication prescribing patterns as a result of an ongoing patient-specific feedback generated from administrative claims
records. This study was a natural experiment in that no comparison group was used and the A&F program was a natural part of care – the study was conducted to examine the effect on a variety of initiatives for the target population (40). Results demonstrated that the intervention was effective at increasing use of under-used appropriate medications and decreased inappropriate medication use.

A second study examined the effect of medication prescribing patterns in Spain for a paediatric population. In this investigation, a quasi-experimental design without a control group was utilized. The A&F program in question focused on providing feedback to primary care providers treating children using antibiotics and anti-asthmatics. Feedback consisted of aggregated results by provider in addition to evidence-based recommendations to optimize dose prescribing. Results demonstrated improvements in the use of antibiotics, but not anti-asthmatics (39). However, it is unclear if performance was maintained as the A&F program only lasted the two years of the intervention.

While the findings from the most recent Cochrane review led to the conclusion that A&F can typically lead to some form of performance improvement – the success of an A&F program is not always guaranteed. For example, a 2004 study attempted to evaluate the effectiveness of a 2-year A&F program on patient reported outcomes for six common chronic conditions. This study was a group-randomized trial in which some primary care teams would receive feedback data while others would not (35). Data were generated from patient submissions of questionnaires for identified health conditions and treatment satisfaction rather than audits of patient charts. Results of the study revealed no differences in outcome between the control and the intervention groups.
Other studies have examined ways in which the effectiveness of A&F may be enhanced. First, a 2013 study attempted to enhance the typical A&F intervention by appending a theory informed worksheet to feedback reports monitoring treatment of diabetes and ischemic heart disease. This study utilized a group-randomized design where the control group received feedback alone and the intervention group received feedback plus the worksheet. The results suggested no difference in outcomes between the two groups. In evaluating the results, the authors attributed their findings to be a function of engagement with the intervention. Given that the worksheet was “passively disseminated” amongst study participants, the enhanced A&F intervention was essentially no different than the methods used in the control group (42).

Enhancement of A&F has also been attempted using more common tools, including financial incentives for performance. A 1999 study investigated the effects of applying financial incentives to A&F to improve delivery of pediatric preventive services. Using a group-randomized design, the investigators conducted an audit of process on three different study groups: A&F alone, A&F plus incentives and a control group. However, results indicated that there were no differences in performance between the three groups (36). Again, the findings of the study were attributed to be a function of engagement with the intervention. Specifically, the authors cited a lack of awareness of both the audit and feedback program in general and the incentives available for eligible participants.

A third study attempted to enhance A&F through the use of academic detailing, practice facilitation and information technology support in the delivery of preventive services. A randomized design was employed which assigned primary care practices to one of two groups: A&F alone (control) or A&F with additional support (intervention). The primary outcomes were the implementation of wellness visits, recall and reminder systems and standing orders. Results
demonstrated that intervention practices implemented more processes than control practices which contributed to improved delivery of preventive care delivery (38). While this study did demonstrate improvement, no statistical tests were conducted on any of the outcome measures. Furthermore, it should be noted that the practices involved in the study were all self-selected, motivated and relatively small primary care practices.

While the above three studies have demonstrated mixed effectiveness of multimodal A&F interventions, a fourth study was identified with much stronger certainty of its findings. This investigation sought to evaluate the effects of a multimodal A&F intervention on colorectal cancer screening in primary care (32). Using a group-randomized design, participating practices were divided into two groups: A&F alone (control) and A&F plus practice site visits and best practice dissemination. Site visits in the intervention group consisted of four half-day visits occurring every six-months throughout the two-year intervention. Additionally, all clinic staff including physicians, nurses and administration participated in the site visits. Results indicated statistically superior performance in the intervention group, compared to the control group (32).

Given the familiar features in study design between this study and the previous three, one must wonder why the results are so drastically different. One reason may be that desired performance change may be a function of engagement with an A&F intervention. This is consistent with the findings of the worksheet enhanced and incentives based A&F interventions – which both cited engagement as an issue which likely affected results (36,42). The colorectal cancer screening study did not suffer from that same problem because the intervention applied was supplemental to an ongoing A&F program for primary care providers in the United States (31,32).
In other words, the 2010 study was more successful because in this context, the use of feedback had already become routinized within workflows of participants. This may be attributable to practice’s participation in the Practice Partner Research Network (PPRNet). The PPRNet is a practice-based primary care research network which collects data from participating practices and providers quarterly feedback reports on specific quality indicators (31). PPRNet has an organizational objective of improving the delivery of primary care services in a broad spectrum of conditions commonly encountered. The A&F program introduced as part of the 2010 study reviewed, was an “enhanced” version of the typical A&F report that practices receive on a regular basis (31,32). Simply put, what makes the 2010 randomized trial unique compared to the other studies reviewed is that the practices were not A&F naïve. Rather, they were all participating in a more mature, developed program, making them much more suitable targets for enhancements to a typical A&F evaluation.

Gaps in the Literature

A 2005 review by Foy et al. concluded that improving the reliability of A&F as an approach to quality improvement can only occur if the features of how and when it works best can be identified (44). This is clearly supported in the primary care context given the mixed effectiveness of the different A&F interventions reviewed. While evidence certainly exists supporting the use of A&F to monitor structure, processes and outcomes – too many studies have applied interventions which appear to end with the project. Further, of the randomized trials reviewed, a common limitation was a lack of engagement or awareness of the interventions. As a result, there is a need for investigations of A&F interventions which are independent of the research programs intended to evaluate them. The PPRNet study is one example of an audit and
feedback intervention which is independent of any individual research project evaluating it (i.e., the intervention itself is not investigator-driven and is not reliant upon research-grant funding).

In addition, little is known regarding the length of time necessary for an audit and feedback cycle to become routinized in any practice context, let alone primary care. Examination and articulation of this factor may assist with the implementation of new A&F interventions in primary care settings which are naïve to the audit cycle. Further, a greater understanding of the factors which contribute to the normalization of an audit and feedback intervention would be a necessary component in determining how and when the intervention works best (44).

A third issue is the length of time during which most studies were conducted. Of the studies reviewed none evaluated an A&F program for more than two years. In addition, no follow-up studies to identify if performance could be maintained were identified in the primary care context. One exception may be a 2011 study which compared the nature and rate of change in primary care performance as a function of its monitoring status (active vs passive) (45). In this case, the study took place in the Veterans Health Administration (VA) which uses a mature audit and feedback system of more than 70 indicators to monitor clinical quality, access, patient satisfaction and value (45). Results indicated that the monitoring status of a performance measure did not significantly impact performance over time (45). The finding that performance could be maintained in the long run is important, however this finding may be attributable to a more mature audit and feedback program. Given that there is no technical limit to a traditional audit cycle, it is still unknown if the gains made from the implementation of a less mature A&F intervention can be maintained. Further, as the A&F intervention is iterative, social and reflective by nature – it would be unwise to employ research methodologies which have a relatively short
timeframe. Therefore, the findings of the studies reviewed may not be reflective of what could be achieved in a “natural” practice context.

Lastly, there appears to be limited use of A&F to monitor or improve performance in more recently introduced primary care practice settings, such as the Family Health Team in Ontario. The one qualitative study reviewed in the FHT setting restricted its sample to physicians (30). This may have been a function of physicians serving as the target audience for the feedback report distributed. It may be of use to explore how other decision makers in primary care engage with, perceive, and respond to feedback in the pursuit of quality improvement. For example, as interdisciplinary practice models continue to grow and develop, the role of Executive Directors and interdisciplinary health professionals will only increase as they oversee the day-to-day operations in the business of primary care. Therefore, the utility of an A&F program to the larger practice as a whole and to other potential recipients of feedback data is an area which requires greater exploration.

Objective

The objective of this study was to evaluate the implementation of an externally developed Audit and Feedback program in team-based primary care practices. This study aimed to address the gaps identified above through an evaluation of an A&F intervention which is relatively young, independent of the research infrastructure, and is targeted towards a newer model of primary care. In order to meet the stated objective, the following research questions were explored:

1. How do participating Family Health Teams implement the Audit and Feedback Program?
2. What factors contribute to the state of implementation described in the first question?
a. What are the barriers and facilitators in deciding to participate?

b. What are the barriers and facilitators to the implementation of the audit and feedback program?

Methods

Basic Research Strategy

A diagrammatic representation of the basic research strategy for this thesis project is presented in Figure 1. The implementation of the audit and feedback program was explored through the collection and analysis of primary data from interviews with key informants within Family Health Teams which participated in the intervention. In addition, quantitative data were used to supplement the qualitative investigation in two ways. First, secondary analysis was used to summarize activity and user behaviour on the web-hosted interactive audit and feedback report. Second, the performance metrics within the Audit and Feedback report were used to provide a descriptive summary of the interviewed sample.

Qualitative analyses in this study were supported using the NVivo software application for Windows (46). In addition, all quantitative analyses were performed using the Harrell Miscellaneous (Hmisc) package of the R statistical software program for Windows (47,48).

This study was approved by the Human Research Ethics Office at the University of Toronto (Appendix 1).
Figure 1. Basic Research Strategy
**Intervention**

Data-2- Decisions (D2D) is a voluntary audit and feedback program targeted towards team-based primary care practices and is administered by the Association of Family Health Teams of Ontario (AFHTO). AFHTO is a not-for-profit association mandated to promote the delivery of high-quality primary health care in Ontario through the promotion of best-practices in primary care and through advocacy (49). The D2D program was initiated as part of AFHTO’s strategic priority to ensure that Family Health Teams are supported to succeed in measuring and improving the quality of care they deliver (50). To date, three versions of the report have been produced: D2D 1.0 in October 2014; D2D 2.0 in July 2015; and D2D 3.0 in February 2016. A diagrammatic representation of the timing of the release of each D2D report along with their respective data inputs are presented in Figure 2.

The D2D feedback report is designed to provide a summary of the participating FHT’s performance on specific quality measures relative to their peers. Each performance indicator included in the D2D report was selected as part of ongoing consultations with an Indicator Working Group which selects measures of interest and defines the parameters for inclusion and reporting. The Indicator Working Group is a panel of leaders from Health Quality Ontario, the primary care research community, as well as representatives from AFHTO-member FHTs. In consultation with the Working Group, AFHTO designed the D2D feedback report to include measures meaningful to FHTs by aligning with the Starfield Model of Primary Care (5,7).

As demonstrated in Figure 2, the timing of the data in each D2D report are not from the same time period. This is attributable to the multi-sourced nature of D2D, which relies on data from four different data sources. First, patient experience data is sourced from specific questions within annual patient experience surveys distributed by FHTs. While the patient experience
survey is not standardized across FHTs in Ontario, a template is available from Health Quality Ontario for FHTs to customize as they see fit. Second, D2D contains measures duplicated from the Primary Care Group Practice Report (PCGPR) available to each FHT from Health Quality Ontario. These measures are summarized based on data managed by the Institute for Clinical Evaluative Sciences (ICES). ICES is a not-for-profit research institute which analyses much of the administrative health-related data in Ontario. ICES is a prescribed entity under Ontario’s Personal Health Information Privacy Act. This means that health information custodians are permitted to disclose their patients’ health information to ICES without consent. Further, ICES may also use personal health information to support the analysis of individual records across a large breadth of data (51).

Third, cancer screening measures are duplicated from data reported in the Screening Activity Report (SAR) from Cancer Care Ontario and the PCGPR for D2D 2.0 and 3.0 respectively. Unlike the PCGPR, the SAR does not use ICES data, but collects data from a range of databases hosted by the Ministry of Health and Long Term Care (52). Fourth, a variety of core and experimental measures are included in D2D which are exported from querying each FHT’s Electronic Medical Records.
**Figure 2.** Timeline of Data and Release of the Data-2-Decisions Audit and Feedback Report
The collection and reporting of data for D2D described above summarize the first two stages of the D2D audit and feedback cycle. At a minimum, FHTs are encouraged to report on the “core” measures of D2D to enable peer-comparison on performance. In D2D, a peer group is defined by the selected setting (urban or rural), teaching status, access to hospital discharge data, and roster size. In addition to the “core” measures, FHTs are able to submit data on expanded measures which are considered voluntary. As of D2D-V3, teams are responsible for the direct entry of data from electronic medical records, the SAR, the PCGPR and patient experience surveys. A screenshot of the direct entry form is presented in Appendix 2. FHTs are required to manually enter the relevant data into their corresponding fields. For data from other feedback reports, such as the PCGPR, the value for the indicator is copied to the corresponding field. Meanwhile, data from EMRs are generated from performing queries on the relevant datasets. In each cycle, teams have approximately two months to enter their respective data, afterwards it is summarized by AFHTO and presented in the D2D interactive feedback report (Figure 3).

Electronic communication is utilized to notify FHTs of the availability of the latest interactive feedback report through the D2D portal. To ensure anonymity of data, each FHT is assigned a randomized access number, known as the “D2D Team Code”, which is used to generate FHT-specific results within the D2D interactive feedback report. Upon receipt of feedback, individual FHTs may do what they please with the information contained within the report. In keeping with the theoretical mechanism of action by which audit and feedback could lead to change, results would be used by FHTs to identify areas of strength and weakness in their respective practices relative to their peers. Following this, FHTs could develop and implement specific action plans to improve performance on desired measures.
Figure 3. Screenshot of Web-Based Interactive Feedback Report for D2D.
While it would not be appropriate to classify D2D as an audit of outcome, D2D does include measures which represent both an audit of structure and an audit of process. In addition to self-reported Practice Characteristics, five domains of quality defined by the IOM may apply to the indicators included in D2D. These five IOM quality domains include: Safe, Effective, Patient-Centred, Timely and Efficient (19). A summary of the core D2D measures used for the most recent D2D report (3.0) grouped by IOM domain are outlined below. In addition, the definition and source for each core indicator is listed in Table 2.

**Practice Characteristics**

Practice characteristics are summarized by six variables. First, practice setting refers to the type of community (rural or urban) which it serves. Second, access to hospital discharge data refers to whether a FHT is able to update its EMR with hospital discharge information. Third, patients served refers to the number of patients served irrespective of their rostered status. Fourth, patients rostered refers to the number of patients formally rostered to the Family Health Team. Fifth, teaching status refers to the practice’s participation in different educational activities. Categories include “academic” (based on formal agreement with a medical school), “teaching” (a self-described status reflecting whether the team hosts a variety of clinical trainees) and “non-teaching” (for FHTs who may host non-clinical, undergraduate or high-school students). Lastly, the Standardized Adjusted Clinical Group (ACG) Morbidity Index (SAMI) score is used as a surrogate measure of the complexity of patients served by the Family Health Team. The index is a set of illness weights for the ACGs using average provincial health care costs, and can be used for examining differential morbidity at a practice level as well as explaining variation among practices (53).
Safe

In D2D, only one measure would qualify as a measure of safe care – Hospital Readmissions. This measure refers to the proportion of patients with an acute inpatient hospital stay for selected conditions who have a subsequent non-elective readmission within 30-days post discharge. Eligible conditions include: cardiac conditions (excluding myocardial infarction), congestive heart failure, chronic obstructive pulmonary disease, pneumonia, diabetes, stroke or gastrointestinal disease.

Effective

Measures which attempt to capture effective primary care include: colorectal cancer screening, cervical cancer screening and childhood immunizations. First, colorectal cancer screening refers to the proportion of patients aged 50 to 74 years old with a fecal occult blood test (FOBT) within 24 months, a flexible sigmoidoscopy within 5 years, or a colonoscopy within 10 years of the report cut-off date. Second, cervical cancer screening refers to the percentage of female patients, aged 21 to 69, who had a papanicolaou smear within the past three years. Third, childhood immunization refers to the percentage of 2-year-old children rostered to the Family Health Team who are up-to-date on their immunization coverage.

Patient-Centred

One measure in D2D is representative of the patient-centered domain in the IOM quality framework – Patient Experience: Involved. This measure attempts to capture the extent to which patients feel involved in decisions about their care or treatment. Data are abstracted from the corresponding question(s) on the annual patient experience surveys.
Timely

D2D includes two measures to capture Timely primary care services: Same day/next day appointments and reasonable wait for appointment. The former is abstracted from patient responses to the corresponding question on the annual Patient Experience Survey based off of the template provided by Health Quality Ontario. The latter, sourced from the same survey, captures patient indications of whether they had waited a reasonable amount of time to get an appointment with their doctor.

Efficient

Three measures from the D2D program attempt to capture efficiency in primary care. These include: Regular Primary Care Provider (Individual), Regular Primary Care Provider (Team) and Cost. The regular primary care provider measure for both team and individual is a summary of the number of primary care visits for a core service that are made either to the physician or team to whom the patient is rostered. Meanwhile, the cost measure represents the total health care system cost for each patient for the specified time period in the D2D report, as calculated by ICES.

Consolidated Quality Measure

In an attempt to capture the total quality of care provided by Family Health Teams participating in D2D, an aggregated quality measure has been developed by AFHTO. This measure represents the weighted-sum of 14-items included in D2D which are intended to reflect patient priorities. Measure selection was informed through the development and administration of a survey by a national patient advocacy group. Respondents (patients) were asked to evaluate the importance of up to forty-five measures commonly used in primary care on the patient provider relationship. Results of the surveys led to the assignment of weights to measures
included in the D2D report, beginning in D2D 2.0. The measure itself was calculated as a weighted sum based on the weights assigned in the surveys. The measure was then refined, using Cronbach’s Alpha methodology, to create a 14-item quality roll-up based on measures which contributed significantly to the measure. At the time of this writing, no peer-reviewed publications exist that validate this measure. As a result, it was excluded from quantitative analysis and was only included in qualitative data if raised by interview participants.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same/Next Day Appointments</td>
<td>HQO Patient Experience Survey</td>
</tr>
<tr>
<td>Proportion of patients who reported they had an appointment the same day or next day after contacting their family health team.</td>
<td></td>
</tr>
<tr>
<td>Involved:</td>
<td>HQO Patient Experience Survey</td>
</tr>
<tr>
<td>Proportion of patients who felt involved in decisions about care and treatment.</td>
<td></td>
</tr>
<tr>
<td>Reasonable Wait for Appointment</td>
<td>HQO Patient Experience Survey</td>
</tr>
<tr>
<td>Proportion of patients who indicate that they could book an appointment within a reasonable time.</td>
<td></td>
</tr>
<tr>
<td>Rostered Visits:</td>
<td>ICES</td>
</tr>
<tr>
<td>Percentage of total primary care visits that are made to the physician with whom the patient is rostered or virtually rostered.</td>
<td></td>
</tr>
<tr>
<td>Hospital Readmissions:</td>
<td>ICES</td>
</tr>
<tr>
<td>Proportion of patients with an acute inpatient hospital stay for selected conditions who have a subsequent non-elective readmission within 30-days post discharge.</td>
<td></td>
</tr>
<tr>
<td>Colorectal Cancer Screening:</td>
<td>ICES</td>
</tr>
<tr>
<td>The proportion of patients aged 50 to 74 years old with a fecal occult blood test (FOBT) within 24 months, a flexible sigmoidoscopy within 5 years, or a colonoscopy within 10 years of the report cut-off date.</td>
<td></td>
</tr>
<tr>
<td>Cervical Cancer Screening:</td>
<td>ICES</td>
</tr>
<tr>
<td>The proportion of female patients, aged 21 to 69, who had a papanicolaou smear within the past three years.</td>
<td></td>
</tr>
<tr>
<td>Immunizations (Rostered 2 Year olds):</td>
<td>FHT-Direct (EMR)</td>
</tr>
<tr>
<td>Proportion of rostered 2-year-old children who are up-to-date for immunization coverage.</td>
<td></td>
</tr>
<tr>
<td>Practice Setting:</td>
<td>FHT-Direct</td>
</tr>
<tr>
<td>Dichotomous variable referring to the type of community served – Urban or Rural.</td>
<td></td>
</tr>
<tr>
<td>Teaching Status:</td>
<td>FHT-Direct</td>
</tr>
<tr>
<td>Proportion of practices who identify themselves as “academic”, “teaching” and “non-teaching” based on their ongoing relationships with medical schools or other educational programs.</td>
<td></td>
</tr>
<tr>
<td>Access to Hospital Discharge Data:</td>
<td>FHT-Direct</td>
</tr>
<tr>
<td>Dichotomous variable referring to the FHTs implementation of a service to update its EMR with hospital discharge information.</td>
<td></td>
</tr>
<tr>
<td>Cost:</td>
<td>ICES</td>
</tr>
<tr>
<td>Refers to the total healthcare system cost of patients rostered/served by the Family Health Team. These costs may include adjustments to reflect age/sex/complexity of patients.</td>
<td></td>
</tr>
<tr>
<td>SAMI Score:</td>
<td>ICES</td>
</tr>
<tr>
<td>A surrogate measure of the complexity of patients served by each family health team.</td>
<td></td>
</tr>
</tbody>
</table>
Theoretical Framework

The Consolidated Framework for Implementation Research (CFIR) is a high-level conceptual framework that encompasses a range of concepts that are applicable to a wide variety of contexts in evaluating the implementation of interventions (54). Introduced by Damschroder et al. in 2009, CFIR was developed in response to the observation that despite using different terminologies and definitions, many of the existing implementation theories overlapped considerably (55). By consolidating the constructs across different implementation theories, Damschroder et al. hypothesized that the identification and understanding of the myriad constructs may be better facilitated in any context if these overlapping constructs were consolidated across the different implementation theories. As a result, CFIR is described as being ‘meta-theoretical’ given that it includes constructs from a synthesis of existing theories in implementation research. However, CFIR does not attempt to describe interrelationships, specific ecological levels or specific hypotheses (55).

In an attempt to capture the whole spectrum in implementation, a literature review was conducted to identify eligible implementation theories and their associated constructs for consolidation (55). The resulting framework is a composition of 39 constructs which are grouped into five domains, which are summarized in Table 3. The first domain, the characteristics of the intervention, includes constructs which are considered to be key attributes of interventions which influence that success of implementation (55,56). Second, the inner setting, refers to the organizational structure, climate and culture in the context of implementation (55). Meanwhile, the outer setting refers to the economic, political and social context within which an organization resides (55). The developers of CFIR caution that there may be considerable interaction or overlap between the inner and outer settings. For example, changes in the outer setting can
influence implementation, which is mediated through changes in the inner setting. Further, what qualifies as inner or outer setting is largely dependent on the context of the implementation environment (55).

The fourth domain describes constructs associated with the individuals involved in the implementation process. Given that organizations are comprised of individuals, implementation is dependent on the actions and behaviours of these people (55,57). Lastly, the fifth domain describes the constructs associated with the implementation process. This domain is based on the observation that successful implementation usually requires an active change process aimed to achieve individual and organizational level use of the intervention (55,58).

The appeal of CFIR is that it offers a platform to explore what works, where, and in what contexts. As the Data-2-Decisions intervention is intended to be implemented within Family Health Teams in Ontario, with no two teams structured the same, CFIR was well suited to meet the objective of this investigation. Another unique feature of CFIR is that it provides additional flexibility in the selection of constructs for the evaluation (55). In this case, researchers can select from the many constructs defined by CFIR based on those which are considered to be most relevant to their study (55).
<table>
<thead>
<tr>
<th>Domain</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of the Intervention(^2)</td>
<td>Factors and features which are considered to be key attributes of the intervention.</td>
</tr>
<tr>
<td>Inner Setting(^2)</td>
<td>Refers to the organizational structure, climate and culture in the context of implementation.</td>
</tr>
<tr>
<td>Outer Setting(^2)</td>
<td>Refers to the economic, political and social context within which an organization resides.</td>
</tr>
<tr>
<td>Characteristics of Individuals(^2)</td>
<td>Refers to constructs which describe the characteristics of the individuals involved in the implementation process.</td>
</tr>
<tr>
<td>Process(^2)</td>
<td>Describes the contracts associated with the necessary steps to implement the intervention.</td>
</tr>
</tbody>
</table>

Notes:
1. Reference: 55
2. Constructs and associated definitions can be found in Appendix 3.
Measures

**D2D Website Analytics**

The D2D Interactive Feedback Report is a website within the members-only section of the AFHTO website. To track website activity of the D2D Report, the research team was granted access to the AFHTO Google Analytics (https://www.google.ca/analytics) account which tracks all activity to the entire AFHTO website. Google Analytics’ reports provide aggregated data on website traffic and user behaviour. In addition to tracking the time spent on a website, a specific page, or the geographical location of a user (limited to municipality), several measures of interest were available for export. These include page views, unique page views, entrances, % exit and the bounce rate. In addition, other analytics terms are summarized below to support interpretation of results.

A *session* is defined as a group of interactions that take place on a website within a given timeframe. A session begins as soon as a user arrives on a website and ends after 30 minutes of inactivity or at midnight. The session is extended every time a user interacts with an element (i.e. opens a new page). Within a session, specific events can also be tracked. An *event* refers to any interaction with the content on a website, such as downloading a PDF file.

Google Analytics also has the capability of tracking activity on specific pages within a website. The most common summary measure is a *page view*. This is an instance of a page being loaded (or reloaded) in a browser. Meanwhile, a *unique page view* represents the number of sessions during which the specified page was viewed at least once. An *entrance* measures the frequency at which users enter the website through a specific page. Similarly, the number of times a user leaves the website through a specific page is referred to as an *exit*. *Percent exit* represents the proportion of exits relative to the number of page views. This indicator is
representative of the frequency at which users leave the website when they view the page of interest. Lastly, the bounce rate is the percentage of single-page visits. These are visits in which the user leaves the site from the entrance page without interacting with the page.

**D2D 3.0 Data**

The data-set used to create the D2D 3.0 feedback report was provided to the research team by AFHTO. Data are summarized by Family Health Team; no patient data were available. The information was used to summarize the practice characteristics and performance on core measures of the D2D feedback report. For the purposes of this investigation, variables were eligible for inclusion they were considered by AFHTO to be core measures of the D2D 3.0 report.

**Qualitative Data**

**Participants**

Family Health Teams were eligible to be invited for an interview if they were participants in at least one D2D audit and feedback cycle and had agreed to participate in an evaluation of D2D by AFHTO. Interview recruitment was facilitated with assistance from the Executive Director of AFHTO (A.H.) and the Executive Director of an Academic Family Health Team in Toronto (K.C.). To recruit specific teams for interview, criterion sampling was utilized to identify eligible Family Health Teams, however the first group of teams were selected by AFHTO – based on their performance on the consolidated quality measure described above. All subsequent groups were selected to maintain a similar ratio of FHTs based on their Setting, Teaching Status, Roster Size, and SAMI score as reported in D2D 3.0.

On the week of February 22nd, 2016 an email was sent from A.H. to the executive directors of the selected family health teams. One week following the email from A.H.
additional email from K.C. was sent to a list of peers encouraging their participation (Appendix 3). Both letters directed participants to register for interviews by providing their contact information using a secure internet-based survey tool (Appendix 4). Two weeks following the initial email, a second group of family health teams were invited to participate in interviews. This process was repeated again two weeks following the invitation to the second group of family health teams.

Participants who completed the registration form were contacted within one business day to schedule their interview. Informants of interest within each Family Health Team were the Executive Directors (ED) as they were considered to be the intended recipient of the D2D feedback report. In some cases, one additional informant may have been recruited to participate in the interview at the request of the Executive Director. To maintain the semi-structured interview format, interviews were limited to two participants.

Criterion sampling of FHTs was determined to be the most appropriate sampling method for several reasons. First, recruitment monitoring was necessary to ensure the interviewed sample was representative of the total eligible sample with respect to roster size, teaching status, setting and patient complexity. Second, given that participation in D2D is voluntary it was necessary to ensure that the broadest pool of FHTs were available for the sample. Third, other sampling strategies, such as stratified purposeful sampling, were not feasible given the limited data available to identify FHTs. For example, although Google Analytics data were available, it is limited in its ability to identify website traffic from a specific family health team. Identifiable information is limited to a user’s city or internet service provider. In addition, the ability to select teams based on their performance on D2D was limited due to confidential safeguards put in
place by AFHTO. Only those FHTs which opted to share their D2D Team Code as part of ongoing AFHTO evaluations were able to be identified.

**Data Collection**

Semi-structured, in-depth interviews were conducted by a single interviewer (D.W.) between March 2016 and April 2016. Each interview was composed of two parts, a discussion about the implementation of D2D in each practice as well as a short think-aloud session examining the usability of the D2D interactive report. Interviews were conducted at the time of the participants’ choosing either by telephone, Skype, or in-person at the offices of their own Family Health Team. Interviews were audio-recorded and a transcription service was engaged to produce verbatim, electronic transcripts for qualitative analysis.

An interview guide (Appendix 5) was prepared, informed by CFIR. The interviews began with an introduction to define the study context and questions to build rapport. It was emphasized that this investigation was independent of AFHTO and that the interest was generally on the topic of Audit and Feedback. Following the introductory stage, open-ended questions were asked to elicit themes regarding how teams use D2D, why they participate in the D2D program, and the resources required to support participation. Areas of interest and issues raised by the participant were investigated in further detail through the use of probing questions. Following the completion of the formal interview, participants were given the option to continue with the usability session, end the interview, or book a follow-up time to conduct usability testing.

Usability testing is a technique used in user-centered interaction design to evaluate a product, such as web sites, by testing it on users. Given that D2D is a web-based A&F program, usability is factor influencing implementation which must not be overlooked. Further, in the
context of the literature on A&F, best practices in usability have been infrequently attended to –
despite the proliferation of web-based A&F interventions. The structure and design of the
usability testing session was based on the guidelines outlined in the book “Rocket Surgery Made
Easy” by Steve Krug (59). Participants were encouraged to “think out loud” as they attempted to
complete specific tasks using the interactive D2D feedback report. For example, participants
were asked to load their team-specific results into the D2D report. As was the case in the formal
interview, probing questions were used to encourage additional feedback from the user to pursue
a specific issue further.

To ensure the interview guide was valid and reliable, pilot testing was conducted in an
interview with an Executive Director of a Family Health Team that did not participate in the
formal investigation (K.C.). Prior to this, drafts of the interview guide were reviewed with a
post-doctoral researcher who had previous experience with qualitative research and was a
colleague of the lead investigator (L.D.). Throughout the drafting and piloting process, revisions
were made to the interview guide to respond to issues such as clarifying questions and
organizing the interview in a sensible manner. Minor revisions were made to the interview guide
on an ongoing basis during the data collection process to respond to issues identified in the
interviews.

Analysis

A framework approach was utilized to analyze the transcripts of semi-structured
interviews conducted with leaders of recruited Family Health Teams (60). The framework
methodology was selected primarily because it had been developed specifically for applied
research in which the objectives of the investigation are set in advance (60). As outlined
previously, the Consolidated Framework for Implementation Research (CFIR) was used as a
theoretical framework for this investigation. Initial constructs from the CFIR framework were selected \textit{a priori} to assist in the design of the interview guide. These themes were subsequently used to populate an initial version of the codebook used for analysis.

The initial codebook used for the analysis of transcripts was a modified version of the sample codebook made available by the CFIR developers (61). To ensure reliability of the coding process, double coding was performed on three transcripts by the lead investigator (D.W.) and one thesis-committee member (J.D.). Each transcript was selected at random, using a random number generator based on the identification numbers of the interviews. The double-coding procedure was executed as follows. First, each coder was responsible for coding two out of the three transcripts selected. Text selected by the reviewer but was not suitable to any of the existing codes were coded as “Other”. The two coders then met in person to review the coding of one transcript in detail. Upon review of “Other” codes, one CFIR sub-construct previously excluded was added to the codebook (Opinion Leaders). In addition, a new non-CFIR code was added to the codebook (Parallel Initiatives). Following this meeting, the codebook was updated by the lead investigator and then both coders re-coded the second transcript. The coding for the second transcript was then validated by D.W. by comparing to the codes applied to the transcript by J.D. This process was replicated for the final transcript, however J.D. performed the validation by comparing her coding of the transcript to the codes by D.W. Following this process, it was determined that the remaining transcripts could be single-coded using the final version of the codebook (Appendix 6).

While an explicit saturation criterion was not used, it was monitored with regular meetings and correspondence between the lead investigator and the supervisor (N.I.). Upon completion of all scheduled interviews, it was felt that saturation had been met. While no
specific interviews were used to seek disconfirming evidence, the interviews themselves were constructed in such a way that the usability testing could challenge assertions participants made during the first part of the interview. Additionally, a single one-on-one interview was held with a QIDSS to get the perspective of the implementation leader for D2D in four FHTs which were not interviewed.

Given the specificity of the research questions, CFIR constructs relevant to the research question were used to construct a framework for thematic analysis. Therefore, a framework matrix was generated for each specific research question whereby the columns represent the relevant CFIR construct and the rows represented each interview. A copy of the framework is available upon request. The specific procedures by which the research questions were answered are listed in the sub-headings below.

**Question 1**

*How do participating Family Health Teams implement the Audit and Feedback Program?*

To answer the first research question, a framework matrix was generated with the following CFIR constructs: *Executing* and *Engaging*. As the question seeks to identify how family health teams implemented the D2D tool, the identified codes were summarized into one of three themes which are representative of the audit and feedback process: Data Submission (Audit), Receipt of Feedback, and Act on Feedback. For each interview, the process by which each element was executed was summarized within the framework matrix. Results were then summarized by examining all three elements individually providing a description of how the family health teams implemented the intervention.
Further, the website activity following the release of D2D 3.0 was summarized under the Receipt of Feedback theme. Descriptive statistics were utilized to summarize page views per day, unique page views per day, entrances, and percent exit. Data summarizing activity between February 1, 2016 and March 14, 2016 were exported from Google Analytics as this six-week period was deemed sufficient to capture typical user behaviour. Website activity could not be compared to the release of the previous feedback report (D2D 2.0) as it was released prior to the launch of the current AFHTO website with Google Analytics enabled.

**Question 2**

*What factors contribute to the state of implementation described in the first question?*

To answer the second research question, two framework matrices were generated to summarize the barriers and facilitators leading Family Health Teams to Participate in D2D as well as barriers and facilitators to implementation of the audit and feedback program. The analytical approach for each sub-question is outlined below.

**Question 2A**

To identify the barriers and facilitators influencing the decision to participate in D2D, a framework matrix was constructed with eight CFIR constructs. Specifically, they include *Relative Advantage, Evidence Strength and Quality, Peer Pressure, External Policy and Incentives, Tension for Change, Leadership Engagement, Opinion Leaders* and *External Change Agent*. Within each construct, the coded text was reviewed and themes were generated to summarize factors which participants suggested served as a barrier or a facilitator in their decision to participate in D2D. Once all the themes had been summarized, they were then assigned to one of three primary CFIR domains: Intervention Characteristics, Outer Setting and Inner Setting. Themes generated from constructs within the “Process” and “Non-CFIR Codes” in
the Codebook (Appendix 6) were re-assigned to the three primary domains listed above based on contextual relevance.

**Question 2B**

To identify the barriers and facilitators influencing a FHT’s implementation of D2D, a framework matrix was constructed with eleven CFIR constructs as well as two non-CFIR constructs. Specifically, these constructs included *Relative Advantage, Adaptability, Relative Priority, Organizational Incentives and Rewards, Compatibility, Leadership Engagement, Available Resources, Access to Knowledge and Information, Opinion Leaders, External Change Agent, Reflecting and Evaluating, Parallel Initiatives and Usability Testing*. Within the framework matrix, themes were generated within each cell to summarize relevant factors impacting a FHT’s implementation of the audit and feedback program. Once all constructs and interviews had been summarized, they were collected into a single list and then assigned to one of three primary CFIR domains: Intervention Characteristics, Outer Setting and Inner Setting.

**Results**

**Participant Characteristics**

Data saturation was reached after 18 interviews. As presented in Figure 4, a total of 118 Family Health Teams were eligible to participate in interviews. Of those, 45 were invited to participate as dictated by the recruitment procedure with 21 expressing an interest in participating in an interview. Three interviews were not completed as participants stopped responding to emails before a meeting time could be scheduled.
**Figure 4.** Summary of Participant Recruitment for Interviews

- **Total Sample in D2D 3.0**
  - $n = 118$

- **Recruited for Interview**
  - $n = 45$

- **Registered for Interview**
  - $n = 21$

- **Unable to Schedule Interview**
  - $n = 3$

- **Interview Completed**
  - $n = 18$
Interviews last a mean of 53 minutes (SD = 8 minutes) ranging from 35 minutes to 63 minutes. Interview formats varied from 1-on-1 to 3-on-1 sessions (Table 4). While the maximum number of participants was intended to be two, one FHT invited a third participant without notice and insisted on their inclusion. Further, participants differed within each interview as different practices had different leaders leading participation in D2D. One interview was held with a Quality Improvement Decision Support Specialist (QIDSS) who is a shared resource among a group of Family Health Teams. Typically, one team acts as the host organization which receives a grant from the Government of Ontario for salary support. AFHTO collaborates with and regularly supports QIDSS, especially with D2D. QIDSS staff receive additional training regarding the processes and procedures related to the audit and feedback program.

<table>
<thead>
<tr>
<th>Table 4. Summary of Interview Format and Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview Format</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>% (n)</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Executive Director</td>
</tr>
<tr>
<td>MD</td>
</tr>
<tr>
<td>QIDSS</td>
</tr>
<tr>
<td>QIDSS</td>
</tr>
<tr>
<td>IHP</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Notes:
1. QIDSS = Quality Improvement Decision Support Specialist
2. IHP = Interdisciplinary Health Professional (Nurse, Dietician, Social Worker, etc.)
A summary of FHT characteristics, comparing the interviewed sample with the total eligible sample, recruited FHTs, and excluded FHTs is presented in Table 5. Comparison of the recruited sample to the total eligible sample demonstrates similar proportions of FHTs based off of setting, access to hospital discharge data, roster size and SAMI. While the standard deviation for the roster size of the total eligible sample is considerably larger, this is due to the inflation by a small number of FHTs which have roster sizes greater than 100,000 patients. These FHTs were not invited for interviews. Comparison of the interviewed sample to the not interviewed sample reveals a less even split in practice characteristics. For example, proportionally less rural FHTs were interviewed compared to those not interviewed. A similar scenario is observed for teaching status, where only two academic FHTs were interviewed. It should be noted however, that roster size and SAMI scores for these two samples were relatively similar.
Table 5. Summary of Core Measures for D2D 3.0

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>118</th>
<th>45</th>
<th>101</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice Characteristics</td>
<td>% (n)</td>
<td>% (n)</td>
<td>% (n)</td>
<td>% (n)</td>
</tr>
<tr>
<td>Setting</td>
<td>Rural</td>
<td>47 (56)</td>
<td>42 (19)</td>
<td>50 (51)</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>53 (62)</td>
<td>58 (26)</td>
<td>50 (50)</td>
</tr>
<tr>
<td>Hosp. Discharge Data</td>
<td>61 (72)</td>
<td>69 (31)</td>
<td>59 (60)</td>
<td>71 (12)</td>
</tr>
<tr>
<td>Teaching Status</td>
<td>Academic</td>
<td>17 (20)</td>
<td>9 (4)</td>
<td>18 (18)</td>
</tr>
<tr>
<td></td>
<td>Non-Teaching</td>
<td>26 (31)</td>
<td>33 (15)</td>
<td>24 (24)</td>
</tr>
<tr>
<td></td>
<td>Teaching</td>
<td>57 (67)</td>
<td>58 (26)</td>
<td>58 (59)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
</tr>
<tr>
<td>Roster Size</td>
<td>20788 (32419.3)</td>
<td>19417 (17383.3)</td>
<td>20994 (34644.8)</td>
<td>19622 (15126.45)</td>
</tr>
<tr>
<td>SAMI</td>
<td>0.99 (0.10)</td>
<td>0.99 (0.09)</td>
<td>0.99 (0.10)</td>
<td>1.00 (0.11)</td>
</tr>
<tr>
<td>Core D2D Indicators</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Patient Involved</td>
<td>89.60 (6.19)</td>
<td>90.92 (5.00)</td>
<td>89.39 (6.45)</td>
<td>91.12 (3.50)</td>
</tr>
<tr>
<td>Courtesy of Staff</td>
<td>89.27 (9.03)</td>
<td>87.15 (9.51)</td>
<td>89.21 (9.5)</td>
<td>89.75 (4.36)</td>
</tr>
<tr>
<td>Cancer Screening</td>
<td>Colorectal</td>
<td>63.51 (10.05)</td>
<td>67.61 (6.38)</td>
<td>63.00 (10.57)</td>
</tr>
<tr>
<td></td>
<td>Cervical</td>
<td>68.97 (9.77)</td>
<td>73.26 (7.16)</td>
<td>68.50 (10.13)</td>
</tr>
<tr>
<td>Child Immunizations</td>
<td>62.13 (20.28)</td>
<td>66.55 (18.59)</td>
<td>61.06 (20.58)</td>
<td>68.93 (17.62)</td>
</tr>
<tr>
<td>Reasonable Wait Time</td>
<td>81.89 (10.68)</td>
<td>82.58 (11.69)</td>
<td>81.80 (10.77)</td>
<td>82.78 (10.75)</td>
</tr>
<tr>
<td>Same/Next Day Appt</td>
<td>54.28 (17.29)</td>
<td>57.05 (15.40)</td>
<td>53.61 (18.00)</td>
<td>59.12 (10.16)</td>
</tr>
<tr>
<td>Hospital Readmission</td>
<td>5.52 (1.33)</td>
<td>5.43 (1.42)</td>
<td>5.56 (1.31)</td>
<td>5.28 (1.48)</td>
</tr>
<tr>
<td>Regular Provider</td>
<td>Individual</td>
<td>64.62 (14.05)</td>
<td>63.67 (11.9)</td>
<td>65.83 (14.1)</td>
</tr>
<tr>
<td></td>
<td>Team</td>
<td>74.19 (12.34)</td>
<td>72.72 (11.91)</td>
<td>75.48 (11.77)</td>
</tr>
</tbody>
</table>

Notes:
1. One interview was held with a QIDSS. As QIDSS work with multiple FHTs (in this case 4) this case was removed from the summarized data for D2D-V3.
Question 1

The following sub-headings include emergent themes which were identified when answering the question: *How do participating family health teams implement the audit and feedback program?*

**Data Submission**

The discussion regarding what data will be submitted is facilitated by the QIDSS (or a FHT’s data specialist) in consultation with FHT leadership. Beyond the standard core measures, the group may discuss what optional measures they may be interested in contributing. Typically, FHTs expressed a willingness to contribute indicators of convenience. These are data which are readily available such as those which are abstracted from other performance feedback initiatives. However, in the event that it is not convenient to submit any data then a FHT will not participate in an audit cycle, as was the case for one FHT in D2D 3.0. In this case, the FHT lacked a staff member with the skills to facilitate data submission. For those teams who did contribute data, several factors were considered in deciding what they would contribute. First, FHT’s will ask their QIDSS to find out (or share) what other FHTs will contribute. Second, FHTs will contribute data to measures they consider to be valid and feasible. As some data require EMR queries, FHTs want to make sure they can successfully export the data and that it is representative of the measure. Some participants explicitly stated that they were less willing to submit data if they were unclear about what was being measured or asked of them. This was cited as a reason teams did not contribute to some measures in D2D 3.0 as they did not understand what they were being asked for, and were not satisfied with the explanation provided by AFHTO when they sought clarification. Third, a FHT’s experience submitting to a voluntary-measure on the previous iteration of the report was cited as an influencing factor in their decision to submit (or not) for
the same measure on the subsequent audit cycle. Repeated submission of a voluntary, non
“core”, measures occurred when data submission was easy. Participants noted that if the data
were too hard to obtain (typically from the EMR), then they would not submit that measure
again. Once a FHT has decided which data they will submit, it is most frequently the
responsibility of the QIDSS to conduct the necessary audits and complete the D2D data
submission form.

**Box 1. Data Submission**

- *She comes to us to say, what do you ... here’s the list of possible things we can submit, which
  ones do you want to, which ones do you not want to. We ask her a few more questions about
details of how this information is gathered, and then she runs it. She sends it to us first, to
say, do these numbers make sense because they don’t always, the searches don’t always work
right, and then we say yeah, go ahead. So right now, it involves ... it was a bit more before
Name-X, when Name-X was submitting it, but now it’s just, it’s a couple of meetings to try to
understand what the measures are, and how to get them.* (ID = 001)
- *Yes, definitely. When I go to the quality committee and say, okay, next round of D2D, we
  need to figure out what indicators we want to participate in, the question is always, well, who
is doing what?.....This one is hard, but if everybody else is doing it, then we should try to do
it, too, or this one isn’t really useful for us. They may or may not be concerned that other
people are doing it. If it’s really not useful, they won’t. But yes, there are a lot of questions
that come about. Who else is participating?* (ID = 003)
- *We try to pull the results from D2D into areas where we need to use similar or those results
with other reporting that’s more mandatory. So things with our quality improvement plan or
with our annual business plan submission or the quarterly submission that we do for the
Ministry. We’ll try to use information from those reports. I think as D2D continues to
mature, we need to access and use the reports more effectively. We’ll start to have
discussions about things when we don’t have to submit mandatory information. Right now
we’re basically using them when we have to and we haven’t used it for our own QI
movements yet.* (ID = 019)

**Receipt of Feedback**

Statistics from the D2D interactive website suggested that FHTs were keen to review
their results for D2D 3.0 (the most recent release) once it was released, however this tapered off
in the six-weeks following. As recorded by Google Analytics, there was a mean of 22.98 (SD =
34.79) page views per day on the D2D Feedback Report Website for the period between
February 1, 2016 and March 14, 2016. In the same period, there was a mean of 14.07 (SD = 18.74) unique page views per day. A scatterplot presenting unique page views per day for the period of interest is presented in Figure 5. Over the six-weeks following the release of D2D 3.0, the interactive report (Figure 3) served as the entrance page on the AFHTO website an average of 0.40 times per session (SD = 0.79). Meanwhile, over the same time period an average of 26.31% (SD = 22.74) of sessions were recorded where D2D served as the last page visited on the AFHTO website.

**Figure 5.** Summary of Website Activity Between February 1 to March 14, 2016.

Concurrent with the release of the D2D interactive report on February 1st, 2016 AFHTO held a webinar to review the results with interested FHTs. While the utility of this exercise was not explored, interviews did elicit information with respect to how FHTs use the interactive report once it is made available.
Typically, the QIDSS staff member will be the first to log-in and review the results for the teams they serve. In a minority of cases, the Executive Director or another FHT leader will log-in and review the results. The QIDSS will often access the interactive report with the ED or FHT leader to review the results with them in-person or via screen sharing using an internet-based video conferencing service.

With respect to the dissemination of the D2D feedback report, interviews revealed that the report is frequently recreated by the QIDSS or the ED. In the latter case, Executive Directors reported that screenshots of the website were copied into a document which is then shared with relevant FHT staff members. Meanwhile, the QIDSS take a more sophisticated approach to modifying the D2D report. Data will be abstracted from the interactive website by the QIDSS into a new sharable format where the QIDSS will add a new comparison where FHTs have the ability to compare the results with the other teams who work with the same individual QIDSS.

These modified D2D reports are then shared/reviewed with FHT leadership such as the executive director, physician lead, or quality lead. The data may also be presented to the quality improvement committee or the board of directors by the QIDSS. However, this observation was not consistent across all interviews.

In the event that data are presented to a board of directors or quality committee, the data are typically presented by the QIDSS. The rationale for this decision was that the QIDSS were best suited to field questions from board or committee members who may be interested in understanding the information in greater detail. Lastly, the extent to which D2D results are reviewed with the staff or clinicians is variable across FHTs. In some cases, only those in a leadership position (such as a board or committee member) will be aware of D2D. In other cases, some components of D2D may be shared at an annual staff retreat or a full-staff meeting.
Box 2. Receipt of Feedback

- The process, in the past, has been that the ED would receive the report, myself, and then I would speak to our chair of the quality committee about the results. We would hopefully present it to the quality committee. The quality committee is made of a physician from each site, and then allied health team members, and site managers. So probably about eight to ten people sit around that committee group, and then that information then is brought to the board at a very high level. (ID = 002)

- QIDSS is typically the one that will let us know, hey, they’re out there if you want to go in and see. She actually also takes the data and … I know we’re going to look at the usability, but one of the things that is a challenge is that the actual dashboards are not … you can’t do one visual. You can’t just see everything in one glance. So, what she has now done is taken all of our results and created our own dashboard of the six of us, which has, I think … when we’re trying to present it to our team, it’s a lot easier to visually present it like that. (ID = 003)

- At this point, what I do is I access the data and I prepare a report back to each of the executive directors and pass it onto them. I know that they have shared it with their board in the past, but it’s basically just been given … it’s just been noted that here’s the D2D report from the QIDSS. There have been no questions back on it, so I don’t think that anybody is actually looking at it yet. (ID = 016)

- We do share the information with our board, however, our board meets once every quarter, so depending on when the information is made available to us or what else is on the board agenda for that day. It might take a back seat to other things we discussed, but people are aware. They’re aware that we are involved with this initiative, with this reporting. We’re still building up to it. (ID = 019)

Act on Feedback

FHTs’ use of D2D results are variable, however no team used this Audit and Feedback tool for the purposes of Quality Improvement. While all teams suggested that they reviewed the data, most were interested in knowing how their performance compared to their peers. In most cases, teams also reported spending a considerable amount of time questioning the validity of the data presented in the report. Specifically, upon comparison of D2D results with internal FHT data some FHTs began to prioritize efforts to improve standardization for data entry into their respective EMRs, as they noted data quality required improvement. However, in the majority of cases this exercise led FHTs to discard that specific D2D result as it was found that internal data were more reliable.
In other cases, a selection of FHTs suggested that their inaction with the results was a function of their unfamiliarity with the audit and feedback process. Stating that they just did not know what to do with the data. Lastly, some FHTs reported that they may use D2D results to support narrative in the preparation of mandatory reports, such as the annual Quality Improvement Plan submitted to the Ministry of Health and Long-Term Care.

**Box 3. Act on Feedback**

- **At this point, I would say no. There are no actual decisions that are being made as a result of the data. There’s some passing interest in it, but there’s not been any actual ... like there have been no quality improvement exercises because of the data yet. But, I foresee that it’s coming.** (ID = 016)
- **We’re discussing it more, but we still kind of ... we get the reports, we look at the reports, we brainstorm them about why the numbers are that way or what can we do to improve them. But we haven’t gone down an avenue yet where we can engage more others or pull more out of (inaudible). We’re still a bit limited on that fact.** (ID = 019)
- **I think that our executive director is obviously the one who is in charge of implementing change. And he’s including me more and more in that stuff too. So, I think ... at this point, we’re not quite sure how to use it. So, we’re just kind of monitoring it for now.** (ID = 011)
- **So, at this point, the level that we participated was much higher than previous, and the report that I produce just became available as of last week. So I think, at this point, it’s more of a learning curve of what do we do next with it.** (ID = 005)
Table 6. Summary of FHT Implementation of D2D by CFIR Domain

<table>
<thead>
<tr>
<th>Theme</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFIR Domain: Process</td>
<td></td>
</tr>
</tbody>
</table>
| Data Submission              | • Led by QIDSS or FHT data specialist.  
• Indicators selected based on convenience.  
• Sometimes FHT select indicators based on what peers submit.  
• Less likely to submit data if unclear about what is being asked of FHTs or if unclear how data will be used. |
| Receipt of Feedback          | • In most cases, QIDSS to log in and review results.  
• D2D feedback report frequently recreated for distribution by QIDSS or Executive Director.  
• QIDSS leads presentation of results to FHT leadership or board.  
• Only in a minority of cases were all FHT staff included in review of feedback. |
| Act on Feedback              | • No team used D2D for quality improvement.  
• Most interested in performance compared to peers.  
• FHTs devoted resources to validate the results in the report, leading to discarding of report.  
• Data may be used for mandatory reporting to ministry. |
Question 2A

The following sub-headings are grouped into themes which emerged when answering the following research question: *What are the barriers and facilitators in deciding to participate in the audit and feedback program?*

**Intervention Characteristics**

*Peer Comparison and Benchmarking*

The inclusion of peer comparison within the D2D feedback report was cited as a facilitator towards participation in the program. Participants reported that this was a unique feature which was not available in other feedback programs such as the HQO Primary Care Group Practice Report (HQO-PCGPR). Having the ability to compare peers provided utility to participants in the form of enabling benchmarking. Respondents felt that the peer comparison would enable them to judge whether or not efforts were required to improve performance of a specific measure.
Perceived Recency

The turnaround time between audit submission (via data entry form) and the receipt of the feedback report was cited as a facilitator to participate in D2D. While this was only cited by a subset of participants, it was revealed that these teams perceived the data in D2D to be more recent, or up to date, than the data in feedback obtained from other sources. As a result, these FHTs felt that participating in D2D would help them plan and allocate resources more effectively.

Box 5. Perceived Recency

- I think some of it was the data that we were getting from the Health Data Branch was so old that we wanted something that was more up to date and D2D was offering the ability to pull the more up to date data. Because Health Data Branch is always a year or two years behind. That really was part of it. (ID = 015)
- I would D2D is on the better end of the reports because it’s current data and you receive it within, it’s usually a month, of when you submitted it, whereas, all the Ministry reports that we receive are usually outdated by a year or two, actually. (ID = 006)
Multi-Source Audit

The fact that D2D is constructed to aggregate specific measures from a variety of data sources served as a facilitator to participate in D2D. This was preferential to FHTs as they could access a single source to view their data instead of reviewing feedback from each of the reports which are used to construct D2D.

Box 6. Multi-Source Audit

- I think one of the main advantages is that you’re uploading and comparing data from a variety of data sources so it’s not just from the EMR. You’re looking at patients seen in seven days, and there’s data from the ICES Primary Care Practice Reports. With those different sources, yeah you’re getting teams to look at the data. (ID = 014)
- Yes, because you have got Health Data Branch information. You have got SAR. You have got our own generated … you have got stuff that we report based on patient satisfaction. So, it’s a nice rounded approach. yes. (ID = 013)
- So, we don’t have any other quality person. So, I don’t have time to be combing through all these reports. So, D2D is the one-stop shop. (ID = 013)

Promise of Future Potential

Participants cited that their participation in D2D was more about the program it will evolve into rather than what it currently offers to Family Health Teams. A selection of participants suggested that the need for a “better data tool” beyond what was currently available. As a result, participants implied that they would continue to participate in order to further the development of D2D to fulfill this “promise”.
**Box 7. Promise of Future Potential**

- It’s just how we’re using it. I think it’s good but I think it should be used as a tool to kind of as a future state of where we want to be going. The work we’re using, the data is not necessarily the number per se. There’s a number of different factors. Are we able to get this information out of our EMR or is this information accurate? What can we glean out of this? So I think it’s helping towards a better future data tool. And I think, as I said, there’s so much complexity and it’s so new for primary care to be measuring, I think it’s getting us where we need to be, not a true comparison of how we’re doing at this moment. (ID = 012)

- And also, it’s evolving. Before, it was like a couple of indicators and now it’s grown ... and there’s a standard process. So, I think we will continue to submit. I think it’s important to support AFHTO’s mission, we understand what they’re trying to do. (ID = 011)

- Yes I do, and I think that that need is right now mostly based on a promise. The promise is what’s going to happen with future iterations, and that it’s going to continue to develop until it actually is a robust, useful, accessible, meaningful exercise. I think we’ve taken initial steps towards that, but we need it to continue in that area. (ID = 014)

**Informed by Participant Feedback**

Another facilitator in deciding to participate in D2D was the unique feature that it is informed by participant feedback. A tension was identified between FHTs being told what to measure by a third party and whether these measures had any meaning or utility for the practices. Family Health Teams therefore saw a need to report on what they thought was valuable by leveraging the infrastructure available through AFHTO.
Informed by Participant Feedback

- I think D2D is more meaningful, because it’s almost a made-by-the-people-for-the-people kind of thing. So, there’s more understanding about the meaning behind what’s being reported, as opposed to somebody that works behind a desk somewhere that has no idea what happens in a family health team. So, I think that, because it’s made by the association that understands what family health teams do and what they’re capable of, that it does have more meaning. (ID = 009)
- Where some of the other external reports that we’ve received they’re based on indicators which we don’t really have a way of impacting the measurement or we don’t know how the measurement is being generated or there could be very little to no importance for us based on that indicator. It’s just something that’s provided to us, where D2D has a member of AFHTO. We have a voice, we have a say of what gets included in that tool. (ID = 019)
- That said, though, D2D has that … unlike many other reports that are out there, it is very much about what the members want and what the members see as valuable. (Interview ID = 003)
- They like the way the D2D is set up, in that the indicators are what is important to family health teams. The feedback, the executive directors were specifically asked what they want to see measured and that’s how D2D evolved, where the QIP indicators were just indicators that the Ministry decided were important, and there’s a lot of contention around some of those indicators, because they’re just weird. But, the D2D ones are solid, they’re more realistic. (ID = 016)

Evidence Base for the Audit and Feedback Program

The evidence base for D2D, performance feedback, or audit and feedback was not cited as either a barrier or facilitator in deciding to participate in the intervention. Further, even though the Starfield Model for Primary Care was used to inform the design and selection of measures to include in D2D – only one team cited it as a factor they considered. A small number of teams cited their perception of HQO’s involvement as an endorsement that D2D was an evidence-based program. On the role of evidence in deciding to participate in D2D, the majority of teams indicated they supported the principle of the intervention (i.e. measurement) and didn’t think a review of the evidence was necessary.
Box 9. Evidence Base for D2D

- I don’t know of any evidence that they put forward to except that they did mention Dr. Starfield as a ... her research project about how she studies those three principles of cost, quality, and capacity. And so, there was that as a framework that they were using that I thought was brilliant as far as going that route, and using that as the principles of where they wanted to go with measuring for primary care ... It helped me sell it within my own self, that we were on the right track. So I’d have to say, I’ve been excited about being involved with it because of those principles. (ID = 002)
- I think mostly the team and I probably accepted the credibility of HQO’s involvement in this in terms of whether or not this was evidence-based or not. (ID = 005)
- I don’t think I have to go to my IT expert and say, do you think measuring how we’re doing is a good idea? It just kind of is. I don’t know how else to say it. I never presented to the group what was the evidence behind D2D. To me, it was like, this is good for QI, this is good for accountability. That makes sense to all of us, let’s go for it. (ID = 017)
- We didn’t go out looking, we didn’t do a lit review to understand why we should be doing this. It was drawn upon our own experiences. Many of the members on our team have been in healthcare for a long time and people are (inaudible), you know, the impression, the understanding that if you can’t measure it you can’t improve it. So that kind of is the belief among those in the family health team. It just made sense. (ID = 019)

Outer Setting

Advise Policy on Direction of QI in Primary Care

A common facilitator to participation is its role in informing the future direction of quality improvement policy in primary care. As highlighted by the tension previously raised between FHTs being told what to measure and the utility of those measures, there is an interest in influencing mandatory reporting to include measures which provide utility to the FHTs. Family Health Teams see D2D as a vehicle to accomplish this, accompanied by advocacy on their behalf by AFHTO to relevant government departments and officials.
**Box 10. Advise Policy on Direction of QI in Primary Care**

- AFHTO has been very engaged and very involved. They have pushed because they knew there was a gap there. And they solved that for our organization, so they need ... That involvement from AFHTO is really, really important. Even if this is something that is taken over by the ministry, I think having them involved as speaking on behalf of the family health teams is really, really important. (ID = 012)

- Getting back to what I was saying before, I’d rather be leading the way than told what to do and how to do it. So I see D2D as our opportunity to really put it out there and say look, to the government, if you’re going to try to measure how well we’re doing and the quality of our healthcare, I’d rather be the one saying this is a known shown evidence-based way to do it. I think we’re working on that with D2D. (ID = 014)

- I need a need for D2D, I think it depends on what the raison d’être of D2D is. My understanding, perhaps incorrect, initially, was that it was clear that there was going to be reporting mechanisms being put in place, being forced upon us from the ministry, and that AFHTO wanted to get in on the ground floor to try to see ... basically to influence the ministry. What’s feasible, what’s reasonable, what’s important to primary care, and to get clinicians involved in trying to influence those decisions. I think from that respect, D2D is important. (ID = 001)

- The other part of it was we wanted to be able to work with AFHTO to start being able to direct where Health Quality Ontario was asking us to go on our quality improvement by using data that was more accurate or more up to date, to create those conversations between AFHTO and HQO. (ID = 015)

- Well, I think it puts us in a good position if indicators that we report on D2D are ones expected of us, say, from the Ministry. I think that puts us in a great position because we’re already able to report on them. (ID = 006)

**Advise Policy on the Value of Family Health Teams**

Participants also suggested that a facilitator to participate in D2D was to support ongoing efforts to quantify and demonstrate the value of FHTs to the government. A tension may exist between measures the government uses to evaluate the practices and those which FHTs themselves think are most reflective of their contribution and performance to the health care system. Further, D2D appears to be one component of a comprehensive strategy to combat bad press and perceived negativity surrounding Family Health Teams.
Peer Influence

While peer influence was not cited as a barrier in deciding to participate in D2D, only a small number of teams cited it as a facilitator. In these cases, participants acknowledged that their awareness of their peers’ participation in D2D served as a catalyst to join the audit and feedback program. However, it should be made clear that these cases were the minority. In the majority of cases, participants challenged the role of peer influence – suggesting they just thought it would be a good idea. Interestingly, some Family Health Teams perceived that they would specifically influence other teams to participate.
Box 12. Peer Influence

- I think we took the plunge regardless early in the first reiteration of D2D because we just thought it was the right thing to do. We didn’t really know what would come out of it exactly so I don’t know how much influence knowing others were participating, but we thought it was safe enough for us to, as I would say, dip your big toe in, see what came back, and how many teams participated. I believe in early days it wasn’t a lot of teams, but I think other teams were convinced to join after the fact when we were able to share some of our results. So maybe we influenced other teams, but I don’t know if other teams influenced us. (ID = 014)

- M: are they supportive of D2D because they know other FHTs are participating? Does that factor in at all? 
  R: No, I don’t think so. (ID = 016)

- M: At the very beginning, you mentioned that you were aware of other organizations participating in D2D.
  R: Yeah, other family health teams.
  M: Did this impact the support for D2D in your setting?
  R: Yes, it did. (ID = 008)

- Yes, other people participating doesn’t really drive our D2D work. It’s great, the more people share, the better idea we have of the benchmarking in terms of generalizing it. But, in terms of working with other people in the area, other organizations like ours, that really doesn’t matter, I don’t think, too much. (ID = 018)

Role as AFHTO Member

A small number of Family Health Teams suggested that their decision to participate in D2D was facilitated by their perception that this was part of their membership responsibilities within AFHTO. However, teams who made this assertion were those with much more sophisticated QI practices compared to the typical FHT interviewed.

Box 13. Role as AFHTO Member

- So I think we participate, really to be good, corporate citizens. I think AFHTO does good work, and if we can help them do that good work, then that’s why we’re doing it. (ID = 001)

- We just want to participate and be an active member for AFHTO. (ID = 007)
Inner Setting

Availability of Implementation Leader

A significant facilitator in deciding to participate in D2D was the availability of an Implementation Leader within the Family Health Team. Except for a few unique cases, participants noted that this role was filled by a Quality Improvement Decision Support Specialist (QIDSS).

Box 14. Availability of Implementation Leader

- Really, it started with the leadership from our Quality Improvement Decision Support Specialists in terms of their working relationship with AFHTO. They were really the cheerleaders for D2D and trying to fan that out across the Family Health Teams that they support, promoting the value of using D2D and encouraging people to contribute data. (ID = 018)
- The executive directors of those nine FHTs, we meet once a month, and these are the sort of things we discuss at those meetings. It really was with ... Our quality decision support specialist, really kind of pushed it too, and he was there to help us get the information. That made it a bit easier. (ID = 015)
- Well, no, at first I was always curious, but we didn’t have anybody. We had a lot of turnover in our staff, so it would have been me, and that’s really not my role. So, it wasn’t until Name-X came on board, our QIDSS specialist, he’s extremely helpful in guiding us and supporting us in his role. So, that’s when we started to ... he’d come up and spend a week with us. So, that’s when we said, okay, if we’re going to take part, we should take part and we have this assistance. He’s kind of guiding us. (ID = 008)

Leadership Role at AFHTO

Another facilitator in deciding to participate in D2D was the participants’ additional roles at AFHTO. For example, five participants disclosed that they were on the AFHTO board of directors at the time of the interview. As a result, their FHT’s participation in D2D may be attributable to the FHT leader’s attempt to support AFHTO initiatives in their other role as a board member. Other participants revealed they were on other AFTHO sub-committees related to D2D. However, the extent to which one’s role on a D2D sub-committee facilitated D2D participation is unclear.
Box 15. Leadership Role at AFHTO

- I think, just as my bias, I am on the AFHTO board. So, to be part of the board, you need to be very, very supportive in the initiatives that come our way. (ID = 003)
- …we actually have some people with our FHT that have helped influence D2D. They’ve been asked to be on some steering committees and stuff. (ID = 017)
- Name-X, just so you know, and you probably do know already. Name-X is on the Board of AFHTO, and Name-X is on the Indicators Working Group. I’m on the Steering Committee, and I’m head of the DM Management Group so we do have a high-level view of this. (ID = 014)
- Well, it doesn’t hurt that I sit on of the indicator working group … And so, I would have to say that we didn’t get a lot of pushback from our board when we suggested starting to get involved with this, because they probably want to support what their ED and lead physicians are involved with. (ID = 002)

**Development of QI Capacity**

A common internal factor among FHTs which facilitated the decision to participate in D2D was the need to develop the capacity for quality improvement. Participants revealed that many FHTs lacked any formal QI processes or committees prior to D2D. Rather than starting from scratch, where they lacked their own expertise to lead QI, many of these FHTs were drawn to D2D and the knowledge base supporting it.
Box 16. Development of QI Capacity

- It’s hard to say because we didn’t really have quality improvement in place before. So, it’s been a great vehicle to advance it. So, at the time, when our organization was old enough to be able to start thinking about quality improvement, that is when it started. So, we could vote as a team. So, we have really grown up with it. So, it’s, maybe, been a consolidating focus. (ID = 013)

- When we started to involve ourselves in quality improvement initiatives in a more formal capacity, we recognized the need to, I guess, retool the organization in terms of training, in terms of resources, to be able to do it properly. We didn’t have the internal expertise to do that. When we were able to hire QIDSS as an improvement decision support specialist that was very helpful, but still having more of a formal process in place that would allow us to structure and tap into data in a more meaningful way was helpful. D2D seemed to provide a method for us to be able to do that, and one that was aligned with more of a provincial initiative as well, so we didn’t feel like we were doing this on our own, but in fact was part of a larger quality improvement community that was engaged. (ID = 005)

- We wanted to participate because we knew the importance of being able to measure the work that we’re doing and we didn’t know how to do this in a meaningful way. We didn’t know if we wanted to ... We thought the D2D would help support us, would help us look at what we are able to measure, what we’re struggling to measure, help us in ways that we’d be able to get that information and be able to compare it with the other family health teams across the province. We wanted to be able to measure how we’re doing, to be able to compare ourselves with other similar groups throughout the province, but knowing that, for us, we were just starting the measurement process. And we wanted to know what we’re able to do and what our limitations were. (ID = 012)
Table 7. Summary of Facilitators in Deciding to Participate in D2D by CFIR Domain

<table>
<thead>
<tr>
<th>Theme</th>
<th>CFIR Domain: Intervention Characteristics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer Comparison and Benchmarking</td>
<td></td>
<td>The ability to compare to peers was cited as a way to evaluate current performance.</td>
</tr>
<tr>
<td>Perceived Recency</td>
<td></td>
<td>Data in D2D were perceived to be more recent than alternative A&amp;F programs.</td>
</tr>
<tr>
<td>Multi-Source Audit</td>
<td></td>
<td>D2D appealed to FHTs as a “one-stop-shop” of data from a variety of sources.</td>
</tr>
<tr>
<td>Promise of Future Potential</td>
<td></td>
<td>Participation was tied to a promise of a better data tool which will come with future development of D2D.</td>
</tr>
<tr>
<td>Informed by Participant Feedback</td>
<td></td>
<td>D2D was perceived to be superior to alternate A&amp;F programs as it incorporated participant feedback in its development.</td>
</tr>
<tr>
<td>Evidence Base for Audit and Feedback</td>
<td></td>
<td>No FHTs considered an evidence base in deciding to participate in D2D.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theme</th>
<th>CFIR Domain: Outer Setting</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advise Policy on Direction of Q1 in Primary Care</td>
<td></td>
<td>FHTs saw D2D as a venue to advocate for their interests in the future direction of quality improvement in primary care in Ontario.</td>
</tr>
<tr>
<td>Advise Policy on the Value of Family Health Teams</td>
<td></td>
<td>D2D was seen as a tool to demonstrate the value and contribution of FHTs to Ontario’s health care system.</td>
</tr>
<tr>
<td>Peer Influence</td>
<td></td>
<td>Knowledge of peer participation in D2D served to influence FHT participation in a minority of cases.</td>
</tr>
<tr>
<td>Role as AFHTO Member</td>
<td></td>
<td>Some FHTs agreed to participate to fulfill their responsibility as a member of AFHTO.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theme</th>
<th>CFIR Domain: Inner Setting</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of Implementation Leader</td>
<td></td>
<td>QIDSS was seen as an essential component which was necessary to participate.</td>
</tr>
<tr>
<td>Leadership Role at AFHTO</td>
<td></td>
<td>Some informants were board members at AFHTO. Participation in D2D was part of their efforts to support AFHTO initiatives.</td>
</tr>
<tr>
<td>Development of QI Capacity</td>
<td></td>
<td>FHTs saw D2D as a way to support the development of capacity for quality improvement in their practices.</td>
</tr>
</tbody>
</table>
Question 2B

The following sub-headings are grouped into themes which emerged when answering the following research question: *What are the barriers and facilitators in deciding to participate in the audit and feedback program?*

**Intervention Characteristics**

*Frequency of Audit and Feedback Cycles*

Participants noted that the frequency of audit cycles for the intervention served as a barrier to implementation. Specifically, they noted that the six-months between the release of D2D 2.0 and D2D 3.0 was insufficient to implement any changes and observe the outcome on the subsequent feedback report. As a result, participants were forced to use the feedback results to inform the extent of their participation for D2D 3.0 rather than completing the audit and feedback process. The value of the feedback to Family Health Teams may be a function of the frequency of audit cycles within a given time frame due to the time cost required to participate in D2D.
Box 17. Frequency of Audit Cycles

- Some of the D2D timelines have been very aggressive. Especially the last one, D2D 3.0, it came out very quickly after D2D 2.0. So, even the opportunity to do the decisions part between the two, there was no time. There was really no time to do it. So, there would have been no time to do any evaluation. All it actually allows you to do is, when the next iteration comes, do a reflection around, what one do we really want to continue to participate in and which ones don’t? (ID = 003)
- We have raised it as a ... in City-X, all the executive directors meet every six weeks and we raised it. Someone was going to take it to the board to say that we felt that the return on investment for the frequency is diminishing rapidly. It’s a very good idea, too many iterations, not well thought out, too ... changing the criteria at the last minute. (ID = 013)
- I thought, for me, it was happening a bit too quickly. I think they could spread it out a little bit more. It felt like D2D 2.0 just happened and then we were getting ready to submit D2D 3.0. And, for us, we need some more time to actually ... Because we’re constantly working on trying to improve our data integrity, trying to get our staff to consistently enter information. (ID = 012)
- I think also tied in with the timeliness is the effort involved. We’re not going to put in this kind of effort on a monthly or every three month basis. Everyone was really relieved when we found out that 4.0 is going to be postponed for three months for lots of good reasons, but that just speaks to it. People are happier to do this every six months than three. That’s a problem, and I think that is going to improve. (ID = 014)

Data

Timing of Measures

The relative timing of the measures included in D2D was cited as a barrier for some FHTs while others may not have considered this issue in much depth. The extent to which participants were familiar with the relative timing of the data contained in the D2D feedback report was inconsistent. In some cases, participants were aware that the multi-source nature of D2D facilitated the presentation of data from different time points. In other cases, participants did not seem to have considered this issue until it was raised in the usability testing portion of the interview (see ID = 003). With respect to the latter scenario, of interest was the assumption that the measures represented a period within months of the release of the feedback report. The ability to discover the relative timing of measures in the D2D feedback report was cited as a further barrier to implementation upon exploration in the usability testing session of the
interview. With respect to the former, participants questioned the value of including data in the feedback report which is more than a year old. Further, only one team identified that ICES data on D2D process measures within D2D 2.0 and 3.0 were identical. Participants suggested that these issues were in conflict with the iterative nature of audit and feedback.

**Box 18. Timing of Measures**

- The problem is, the last I heard, the data that is being given to us for the 30-day readmission is a year too old. So, I don’t understand how you can analyze ... So, a perfect system would be real time data. (ID = 017)

- You know, it’s old data. It’s like a newspaper that’s a year old, and picking it up, and reading it. (ID = 002)

- Where would you find information about the dates that the data was collected for each indicator?
  R: I have no idea.
  M: Is that something that you think would be valuable?
  R: Yes, probably. We know the date of the report because it is here, February 1st, 2016. So, you assume the data was collected within a few months prior to that. But yes, it would probably be helpful to have it somewhere in that header. (ID = 003)

- R: Yes, but see, the data dictionary definition gives you within the last three years, but it’s not telling me specific date ranges. I run it on, say I run it on, the last time was December, so I would’ve ran it in December, but other teams could’ve ran it beginning of December, later than I.
  M: Do you think, though, that having and presenting that date information is important or would be useful in interpreting the report?
  R: Well, especially if you want to compare to other teams. That definitive date range would probably be more comparative than, well, I ran mine a month after you and so I have now more patients who have been screened this year, and that type of thing. (ID = 006)

- I have another disadvantage, is some of the indicators that are on D2D are reports we get from HQO. So, we submitted in the fall for, what was called ICES then, new ICES data, but then they just sent the old stuff from the previous time we did D2D, so we were really reporting the same thing as we did the last instance, which really isn’t that beneficial because your numbers are the same. So that’s just a ... It’s not really AFHTO’s fault because HQO is the one prepares that data, but it doesn’t really make any sense. (ID = 006)

**Sensitivity**

The sensitivity of an indicator to capture team performance was cited as a barrier to implementation. At issue was the observation that many indicators are representative of
physician behaviour, rather than team behaviour as the data are based off of billing codes. In these cases, many FHTs noted that some indicators, such as 7-day follow-up were tracked internally as it is a task which may be completed by staff other than physicians. Given a FHT’s ability to track these measures internally, and the conflict in the D2D indicator’s sensitivity, FHTs have less incentive to pay attention to the specific result in D2D. Participants noted that the inclusion of these types of measures presented a challenge for the legitimacy of D2D among FHT staff.
Box 19. Sensitivity

- But how they’re tracking it is not reflective of the spreadsheet that we’ve done and internally tracked our process. And I think some is made of billing codes. Well, only physicians are billing. The billing is different if you’re seeing your patients are coming from long-term care because the codes are different. So our numbers are really not reflective of internally what we’re measuring. (ID = 012)
- But people, sometimes, don’t believe the data because it’s wrong, but it’s not wrong for that reason. It’s wrong because we’re looking only at physician billing. (ID = 003)
- I think, unfortunately, D2D relies heavily on billing data at this point, reported from ICES. So as to, you know, it can get really deep into, does that really reflect the team when our nurse practitioners are billing. I think that’s the struggle that we have with some of that data. So, unfortunately, we usually rely more on EMR data, which really reflects what the FHT is doing instead of an individual physician. (ID = 005)
- Just sometimes the frustration of the data. Again, we’ve had people saying well this … So it’s hard to say, when they look at the data, this can’t be right. I’ve had feedback on all this discounting and distrustful of the performance (inaudible) itself. But we’ve discussed it, I’ve emphasized several times through our discussions we need to be fair. This is just the beginning of trying to figure out how, as a primary care centre, we are able to measure our performance and the value that we bring to our patients. So I think it’s always reframing that back but I think that has been … When there’s feedback or data that we receive that we don’t feel is reflective of what we’re doing, there is that backlash a bit of this is not really working, this is not true. So it kind of discounts that whole measurement piece. (ID = 012)
- My one physician and I, and I think he has a very valid point. They’re telling us they want to know what percentage of patients can get same day appointments or can get an appointment with their family physician. So, we scored low, but we have IHPs that work with these family physicians and so they don’t need to go see their doctor if they want to have their blood pressure checked. They can see a nurse. That’s why family health teams were created. But, no, that doesn’t count. That doesn’t count because it’s not their family physician. We all have a very difficult time with that mindset. We’re a family health team. We work as a team. We work in conjunction. If that person that has to have their blood pressure checked, if the nurse sees that there’s an issue, she goes to the physician and then it’s discussed or the doctor walks in. So, we do work as a team that way. Patients can come in and see a nurse. Like today if you need anything, it’s done today. We have a walk-in, they can see a nurse. And, if that patient then needs to see a doctor, then they go on to the doctor. But, to say, oh, Name-X, Joe Blow is your doctor and you want to see Dr. Joe Blow today, I’m sorry but I can let you see a nurse for this. So, I think that that’s where we struggle with the focus of D2D. (ID = 008)

Duplication

Duplication of data from other feedback initiatives was cited as another barrier to implementation. Participants cited confusion with respect to duplication of measures as it
prevented D2D from fulfilling the promise of being more up to date than competing performance feedback initiatives. However, one team suggested that the duplication of data served as a facilitator to implementation. In this case, the enhancement via peer-comparison enabled the establishment of performance targets on some indicators.

**Box 20. Duplication**

- I’m not sure why they are in D2D if they are already in the QIP because, again, some of it is based on data that isn’t timely. (ID = 003)
- But, there does need to be some work on AFHTO, and I find there’s a little bit of duplication in the system for us. (ID = 007)
- Some of the indicators are similar to what we’re inputting into our quality improvement plans and what we’re putting into our annual reports and quarterly reports up to the ministry. It does help in the sense it gives us more context, that there is some comparison to other groups. Because, as I mentioned, we only have one-way feedback, so this has helped us around, for example, with our quality improvement plan, knowing from D2D where we’re at with other groups. Our target was for each individual to be at least at that average point. So that helps, it helps with our planning process because it gives us a target to work towards. (ID = 012)

**Definitions**

Another barrier to implementation was the lack of transparency regarding the way in which the data are defined in D2D. Specifically, participants raised methodological concerns regarding the ambiguity, accessibility, and reliability which limited their ability to trust the feedback report. Methodological ambiguity led some FHTs to be surprised that their data were being used to calculate measures which they did not think they submitted data to. Further complicating this matter, participants expressed frustration that the support material (data dictionary and support videos) failed to clarify their concerns. Many did not understand why the complete methodologies were not easily accessible. Lastly, some FHTs were concerned with the methodological consistency of data queried from peer electronic medical records. Participants
suggested that the peer comparison feature could not be fully utilized without ensuring methodological consistency across sites.

**Box 21. Definitions**

- I don’t think, to be quite honest, there’s a lot of trust in D2D, because we are having results that we don’t know where they’re coming from. That would be like you putting in a question and then you get six versions of the question back, and it’s like, what does that mean? (ID = 007)

- In other cases, my sense is there’s a lack of confidence in the reliability of the numbers in terms of comparisons, because we don’t believe that people are measuring things the same way. So, the same-day, next-day access data set is a good example of that. How that’s actually being tracked and measured, I don’t think it’s standard, I don’t think it’s reliable. So when I look at the comparisons, I don’t find them helpful, I don’t have a lot of confidence in the validity of them. (ID = 005)

- It’s often unclear, that’s part of what the trepidation is. I’ve had the conversation about this particular one, time spent and I’ve asked things like, I don’t know that, whatever your methodology looks like, I don’t know that you’re factoring in patient complexity, in terms of how much time providers need to spend with people, based on their complexities or their patient profile. And people will say, it’s in there, but they can’t show me how it’s in there. And so then, just saying that it’s in there doesn’t mean that I’m going to trust that it’s in there. (ID = 018)

- There was one on capacity that I really didn’t agree with whatsoever because we didn’t like how it was framed. And then, when we looked at the output in D2D and we saw that graph. And I was listening to a couple of YouTube videos and it was part of the meeting to just explain what that graph meant. And even they said that they didn’t necessarily understand what that graph was. So, why are we contributing to it? So, I think, really, what is being asked of us, there needs to be a reason behind it?... but I find that D2D hasn’t been afraid to say we’re still working out this particular indicator, bear with us, it’s early stages. And I think the capacity is an excellent example. The problem is, I have an issue with, if you didn’t get it right the first time, in terms of what you wanted to measure, I’m okay with that, that’s research, that happens. But then, maybe come out and admit that you’ve got to go back to the drawing board. (ID = 017)

- There’s no real definition on what goes into those graphs. So if we’re looking at the data quality or the quality roll-up indicator, the average person wouldn’t know the 14 indicators that go into it. We noticed that the quality indicator changes every time you go into it, so even that makes it a little difficult. So one time you go into it, it’s one number, and then it changes the next time based on what you reported on, so I think that threw a lot of the QIDSS specialists in a little bit of a loop. We didn’t realize that was going to be an issue. I would not send the link out to someone to go to. I would take on the role of sorting that out myself and then providing the information afterwards. (ID = 005)
Peer Comparison

Participants cited the clarity of the peer comparison as another barrier to implementation. While participants are supportive of the principle of peer comparison, many suggested that their inability to judge the validity of the comparison, beyond practice setting, teaching status and roster size impacted the utility of the comparison. Two primary concerns emerged with respect to peer comparison from the interviews. First, the relative sample size used to calculate a statistic should be considered as a factor which may influence peer comparison. Second, the extent to which patient complexity (as calculated by SAMI) is used to adjust peer comparison values should be disclosed.

Box 22. Peer Comparison

- The thing that I’m always pressing people for is how the SAMI score factors into adjusting different rates. So, if that was even more explicit. (ID = 018)
- I would just say how many actual values were submitted? It’s helpful. Because, for some of these I think that, especially for the patient feedback, we might have a smaller sample size than some of the others, which isn’t necessarily a bad thing. But it would be helpful to know. (ID = 011)
- I think, for us, the one caution in my mind with looking at it is that I don’t think, in terms of looking at your peers, they give you things like academic teaching, non-teaching, setting, urban/rural but there is no option to pick your model type. So, we’re a community urban Family Health Team, operating in a blended salary model with our doctors. I would like to be able to compare, like that’s my fear. Not all family health teams or urban family teams are family health teams that operate in the same model, because they’re different. The focus is different, the way overhead is calculated is different, everything is very different. (ID = 018)

Website Design

The design of the website which hosts the interactive feedback report was cited as a barrier to implementation for Family Health Teams. Several factors were cited which limited the usability of the website and assist users to interpret their results. First, the inability to access the data dictionary within the feedback report was cited as problematic. At minimum, participants
expected to see a direct link to the data dictionary, however some expressed a desire for a more sophisticated website where they could interact with a specific indicator label to see the definition. As users felt the website was difficult to navigate, even after its redesign in the Fall of 2015 some participants saved copies of the data dictionary to their computer. In the event that the data dictionary was ever updated, these FHTs would not have access to the resource. Second, participants noted the presentation of the data in the feedback report may add unnecessary complexity, challenging interpretation of the data. For example, one participant suggested that the presentation of the graphs was in conflict with convention with respect to the labelling of axes and scales. Third, participants expressed frustration at the limited functionality of the feedback report. The inability to export the data, share the results or simply printing the results were all cited as barriers to implementation.
Box 23. Website Design

- There’s no option to, maybe, export it out of the web site into Excel or anything. The graphs, you know, to get to the number you had to move your cursor over so that meant me sitting there writing out all the numbers on a piece of paper and then transferring it to Excel. The expanded data wasn’t on the web site at all, so that was a little bit disappointing, so that meant a lot of work afterwards. And then, I guess, when you went to go click on the targets it brought up a PDF. There was just some disconnect with the whole report on how it come out. So if there was any way to export it so that you could get a two or three page report that you could hand to someone, I think that would be a lot more helpful than a lot of links that didn’t really seem to connect well. (ID = 005)
- The presentation of D2D could be better, like the way that the website ... the logic behind the way that ... like you drill down and stuff into that is different. Most of the other reports are in some sort of chart format, where it’s actually just figures. There are no graphs and stuff like that. So, D2D attempted to visualize a lot of that data, and by doing so, makes it sometimes awkward to understand, which is kind of the reason that I make my own report from it. (ID = 016)
- M: Where would you look for the data dictionary?
  R: But for me to look for anything on the site, it’s easier for me to go to my computer because the site is not easy to navigate. (ID = 003)
- I would like to see, I guess it’s just the way that I’ve always learned, is when you have a graph there’s typically a title for the graph and then information. So like effectiveness which is the area where the child, the immunization information is under. It doesn’t really tell me effectiveness of what. It just seems to be kind of out there. (ID = 019)
- I would have to go into another page, click on colorectal screening, it would take me to the AFHTO page that would tell me where to go for my information. And it usually takes me to HQO or something, some long report, or actually an AFHTO page. You have to drill, drill, drill down. It would be great if you could hover over this and it would say, this is this. It’s a lot for work to be able to figure that out. Like, I can’t remember whether it’s over 50, or over 65, or over this, that, you know, the population and the exact, you know, numbers that you need to be able to put that in off the top of my head. So if I was presenting it to a group I would do all that pre-work ahead of time, I would go and print those pages, or make sure it’s off the top of my head. (ID = 002)

Outer Setting

QIDSS Dependence

The availability of the Quality Improvement Decision Support Specialist (QIDSS) served as a facilitator to a FHT’s implementation of D2D. While there were a select number of cases where the QIDSS was not needed, this was limited to FHTs with in-house expertise to perform
tasks which are typically QIDSS’ responsibilities. In all other cases, a FHT’s participation in D2D is dependent on the availability of the QIDSS resource.

The QIDSS serves as the implementation leader for D2D at most FHTs. Their responsibilities include the organization of data collection, data submission, accessing the feedback report, and even the dissemination of the results to the FHTs they work with. Unfortunately, as QIDSS are a shared resource among a group of FHTs, they are often limited in their ability to support implementation beyond the audit and receipt of the feedback report. Further, FHTs argue that given that the QIDSS time is shared they mostly end up helping with D2D alone, due to the amount of time it takes to participate, and as a result limits the QIDSS’ ability to contribute to other quality improvement activities at the FHT.

The extent to which FHTs are dependent on the QIDSS to participate in D2D is quite significant. This is best exemplified by ID = 009 which lost their QIDSS for a short while and ended up failing to participate in D2D 3.0. This dependence was corroborated when raised in an interview with a QIDSS alone. While the QIDSS suggested that some may participate a little bit, teams with little interest in D2D would not continue to participate.
Given the complex landscape of performance feedback initiatives in which teams can participate, the relative priority of D2D may serve as both a barrier and a facilitator to implementation depending on participation in other programs. Interviews with participants...
revealed varying participation in a wide array of audit and feedback activities by Family Health Teams. At a minimum, all teams participated in mandatory programs such as the annual Quality Improvement Plan (QIP) submitted to Health Quality Ontario (HQO) as well as quarterly Schedule A submissions to the Ministry of Health and Long Term Care (MoHLTC). Additional programs which all teams participated included the Primary Care Group Practice Report (PCGPR) from HQO and the Screening Activity Report (SAR) from Cancer Care Ontario (CCO). Lastly, all participants noted that they collected patient experience surveys.

In addition to the reports listed above, most participants revealed that they had their own internally administered audit and feedback programs with varying degrees of sophistication. Additionally, a selection of teams revealed they received performance feedback reports through programs such as Electronic Medical Record Administrative Data Linked Database (EMRALD) from ICES and the Canadian Primary Care Sentinel Surveillance Network (CPCSSN) from the College of Family Physicians of Canada (62,63). Additionally, some teams participated in regional initiatives such as the Normal Simcoe Community Health Information Portal (NSCHIP) or the Ottawa Model for Smoking Cessation Program (OMSC).

Analysis of parallel performance feedback initiatives in concert with the relative priority of D2D revealed several factors which may contribute to the relative priority of a performance feedback report. These include data quality, data timeliness, the ability to link to patient data and financial incentives. D2D often ranked lower than other performance feedback initiatives due to varying combinations of the factors listed above. A consistent observation was that these factors were directly linked to the perceived value of the feedback report to the practice. With respect to financial incentives, the SAR by CCO was frequently mentioned as a tool used as practices attempted to maximize their cancer screening bonus. Further, initiatives such as the QIP and
Schedule A (submitted to HQO and the MoHLTC, respectively) were also of higher priority due to regulatory requirements linked to FHT funding. As a result, when D2D was released close to QIP reporting time, as was the case for D2D 3.0, it often took a back seat at Quality Improvement Committees and Board Meetings which were focused on readying their respective QIPs.

Lastly, the maturity of a Family Health Team’s quality improvement efforts may have also contributed to the relative priority of D2D at participating FHTs. For example, D2D was observed to be of higher priority at Family Health Teams which lacked internal reporting (or sophisticated internal reporting) or those which did not disclose their participation in external programs such as EMRALD or CPCSSN. Teams with sophisticated regular internal reporting, which also participated in external programs such as EMRALD and CPCSSN noted they participated in D2D to help out, not because they saw a need for their own practice.
Box 25. Relative Priority

- We share it. We make reference in terms of what it is we’re looking at. We don’t print the whole report and specifically just print the report to talk about the report. It’s more of a source that we use to do our jobs, not focus our jobs around D2D. (ID = 018)
- We rely on our patient surveys and our diabetes stats quite highly, because they are monthly, so they’re real time, and because with our diabetes stats we can then drill down into those stats and find out who the ones are that we’re missing. So that’s very concrete for us. Some of the HQO and the D2D stuff is more higher level. (ID = 015)
- I think probably D2D more than the others but I think, with our quality improvement plans, it’s the internal stuff that we’re measuring that I find I rely on the most. I think it’s telling us the truest pictures. (ID = 012)
- the Ministry of Health operational planning things, just because it’s a pretty big endeavour that takes a lot of time to input and requires every single person’s input that’s responsible for providing some kind of service or programming. And it’s more closely tied with funding, so I think there’s a little bit more emphasis placed on that. (ID = 009)
- other problem is we have all of our information, which is second to none through Simpson and UTOPIAN. We now have a shared data. We’re housed with the hospital, where we can look at acute and primary and see how we’re going in that area. This is a bit of a make work project for us, but we participate because we thought we’d be very useful to be part of the bigger picture. (ID = 007)
- I think what I meant to let you know is that we try to speak to the quality committee this time for 3.0 prior to the board meeting, but it didn’t quite make it through the agenda. And so, the board got presented first, and hopefully we’ll have time to talk to the quality committee about it next time. But our priority was the QIP that’s due out on April 1st, and we had a lot of discussions around that as a priority as opposed to the dissemination of these results. (ID = 002)
- I think that we’ve spent quite a lot more time and got more value out of the Cancer Care Ontario SAR Report than anything else. (ID = 014)
- how many charts, if we randomly audited X number of charts, how many charts would identify a patient’s smoking status. And they might say to everybody, at some point in the next month, we’re going to audit the charts. So, do what you’ve got to do with them. So then, two weeks would go by and they would audit and they would let every provider know what percentage of their charts had smoking status or allergies or whatever. And that would be a nice internal review for us that would bring to light certain deficiencies. Whether it was proper documentation in the chart, it might be what percentage of people over 65 have been immunized against pneumonia? And that had nothing to do with D2D. (ID = 017)

Available Resources

Time

Time may serve as both a barrier and a facilitator to implementation of D2D. Specifically, time is an issue in two different contexts of D2D. First, participants were in
agreement that the data submission process for D2D is time consuming. This is especially the case when comparing to other external audit and feedback programs such as the PCGPR, the SAR, EMRALD, or CPCSSN. This serves as a barrier to implementation as teams are forced to consider the amount of time it would take to contribute data rather than selecting measures which they deem to be of interest or of strategic priority. This time effect is compounded when the timing of D2D iterations are close together, as was the case for D2D 2.0 and D2D 3.0. Some teams expressed that they lack the resources to contribute at a high frequency of audit cycles in concert with other reporting and clinical responsibilities.

Second, time was also a concern with respect to a FHT’s ability to act on the feedback or enable discussions surrounding quality improvement. As highlighted above, the FHT’s are dependent on the QIDSS to facilitate implementation of most stages of the D2D feedback report. In interviews, QIDSS themselves revealed that they lacked the available time to engage on anything beyond data submission, preparation and presentation of results. Further, FHTs also expressed an interest in the QIDSS spending more time on the development of quality improvement programs, beyond their current responsibilities. In addition, FHTs reported that they struggled at times to balance the development of quality improvement initiatives without compromising the availability of the clinical services. Neither of these issues were reported among teams with greater resources, such as those classified as Academic or at FHTs with their own data specialist positions who are full time staff.
The other thing I like about EMRALD Sapphire, and the Primary Care Practice Report is that they’re generated automatically, I don’t need to … my QIDSS person spends a lot of time giving D2D data. (ID = 001)

Some of the other systems we just send them the data, and they send us the reports. With D2D there’s more work involved, and we’re trying to minimize that, but there certainly is more work involved. (ID = 014)

If they keep doing it at the rate that it’s going right now, I’ll probably not participate. I’ll probably talk to our group about not participating and having our quality improvement data support specialist do something else, because that is really all she has been doing, is getting it ready and doing these submissions. So, if this continues, I would be making it not a priority. (ID = 013)

I don’t think that, from a team perspective, the teams have the time and capacity to attend to this to the level that perhaps HQO thinks we should. So it’s one thing to say you have the staff resources in a QIDSS specialist like QIDSS to assist, and it’s wonderful. But when you look at how people are spending their time in organizations, and you ask them to engage in D2D or other quality improvement initiatives, it takes time, people attend meetings, their tracking data, (inaudible) there’s a whole bunch of pieces. When they’re doing that, they’re not seeing patients. And so, (inaudible) the quality improvement initiative is, the Ministry going back to your thing around the policy climate, is really trying to drive increased accessibility. When I have 10 people in a meeting for two hours talk about quality improvement, that means they’re not seeing patients, which reduces access of our patients to our team. So, in terms of resources, there needs to be greater organizational capacity to be able to plan, develop, and implement quality improvement initiatives beyond a part-time QIDSS specialist. Otherwise, we really are working in conflict in terms of trying to give patients greater access to our providers, while, at the same time, distracting our providers by trying to engage them in things like quality improvement initiatives or other things that seem to come down from the Ministry within that particular policy climate. (ID = 005)

Because it’s difficult to extract information from the EMR, our HPs and our RNs have to spend a lot of time extracting this data. I think there is value in it because I think we have to demonstrate … so, it’s just that it’s hard to pull the data and we don’t have a quality improvement person. So, it’s taking time away from patient care. (ID = 013)

I think now as a QIDSS, being a resource myself, I think it’s enough that I’m able to collect the data and submit on it, but being able to act and implement those changes that are required to lead quality improvement, there definitely needs to be some more resources put in place, especially if it continues to grow. One person can only do so much with the time so the face time I have at each FHT and the influence I have, there really needs to more of us I would say. (ID = 014)

### Human Resources

A consensus among participants was that a certain level of human resources was necessary to facilitate the implementation of D2D. However, this was not a problem for larger
FHTs or those with specialized resources such as data managers. In these cases, the FHTs have access to a full-time employee whose responsibilities include data management, analysis and the facilitation of quality improvement. In a few cases, some FHTs reported that they had the budget allocated to fill this type of position, however they faced challenges with respect to recruitment of a qualified individual from their local community.

The most typical scenario was that the responsibilities of such a specialist were distributed among a mixed variety of full time and part time staff at the FHT. To varying degrees these responsibilities are typically handled by a committee of staff. However, this exposed FHTs to vulnerabilities as staff turnover impacted their ability to implement D2D. In the context of all the additional reporting FHTs are required to do, some FHTs speculated that without a full-time resource at their disposal, their ability to implement any performance feedback program would be limited.
Box 27. Human Resources

- And, we have seen with D2D, we were usually at the top. But, we also have a lot of resources to help us get to that top. So, some of our comparators do not have any data managers gleaning their data. They don’t have all of that. (ID = 007)
- R1: Our data person is a dietician.
  R2: Yes. She is a half-time dietician, half-time data person.
  R1: She was never a data person. We just gave her the job and she learned it on the job, which was great. I think that is also one of the concerns, too. From a knowledge transfer perspective, if she left tomorrow, we would be in a lot of trouble. (ID = 003)
- I did have somebody that was in that role, like a program manager person that was very good at statistics, but I don’t have that individual any more. And, if I could recruit one and hire one, that would be great. And, the funding is there for me to do it, it’s just that that’s not who has filled the role. I think the support is there for me. I can’t speak to other family health teams. I think it’s available to me and I just can’t find a resource in my community. (ID = 008)
- There has to be to be able to get this information, to be able to track it over time, but we’ve utilized summer students. We’ve utilized student placements, a partnership we have with the University-X, and Cancer Care Ontario, to help us with our preventative care information. So we’re utilizing as many resources as we can and being very innovative and creative to do that because we feel it’s really important but it definitely has been a challenge. (ID = 008)
- So, this year, we are only doing the bare minimum. So, what we are focusing on is what is required by Health Quality Ontario, which informs D2D. That is it because there is just too much turnover this year. So, that is what we decided to do. (ID = 013)
- I absolutely think that those resources need to grow as well. I’ll use the example, and I don’t want to put QIDSS in a position, that she’s employed by our team but supports seven teams. That may not, depending on how detailed, and it’s not just data extraction. It’s the improvement in analytical work that needs to be done. I don’t know how an individual will be able to maintain that workload over a period of time, unless the metrics stay very small and narrow. But if we truly want to dig down, we’re going to need people, resources, to do some of that work. Most teams aren’t structured the way we are, but other teams, I’m just speaking in general, I don’t think there are enough resources to do the quality of work that will be required moving forward. We’re doing it now to the best of our ability, and we could still do better if we had more time to do that. It’s just one of many things we do. (ID = 014)
- I think right now it is a bit of a data overload and reporting overload with the capacity of the teams. (ID = 014)

FHT-FHO Relationship

One theme which emerged from interviews was the impact of the relationship between the Family Health Team (FHT) and the Family Health Organization (FHO) on the implementation of quality improvement initiatives. In the case of D2D, discussions revealed that
the FHT-FHO relationship served as both a barrier and a facilitator, depending on the specific setting.

The relationship between the FHT and the FHO varied among the settings which participated in this investigation. In some instances, participants suggested that the FHT and FHO worked as partners and have integrated themselves to support each other in pursuit of shared organizational objectives. At the other end of the spectrum, participants described a situation where members of the FHO viewed the FHT-FHO relationship as two independent entities and were disinterested in team objectives. Given that D2D is comprised of a variety of physician specific measures, a FHT’s ability to act on feedback or attempt to improve data quality related to physician specific indicators is dependent on the strength of the relationship between the two entities.

The suggestion that the QIDSS job would be more successful as an employee of the FHO instead of the FHT is perhaps the most revealing example of the impact the FHT-FHO relationship has on a FHT’s ability to implement D2D. Discussions with participants revealed several factors which may affect a FHO group’s, or individual physician, willingness to cooperate on quality improvement activities. First, physicians who are close to retirement are not interested in doing more to modify their EMR practices. Second, physicians who have responsibilities in other care settings such as the emergency department at a community hospital lack the time to invest in QI. Third, some individual physicians may see the FHT as an entity to which they are not accountable. As a result, FHTs have to carefully manage their relationships with the FHO and even individual physicians. As demonstrated in some of the quotes below, they rely on various engagement strategies as well as patience and persistence.
**Box 28. FHT-FHO Relationship**

- Yeah, because the team is very separate from what the actual organisation is. The organisation is made up of two teams, the FHT team and the FHO team, and the FHT team is very, very separate from the FHO team. If I could go back in time, I would try to figure out a way to set up the structure that wasn’t an employee of the FHT, that I was an employee of the FHO, and that would be my angle, you know what I mean, that I actually work for the doctors? (ID = 016)

- I think with our providers being an independent FHO, they don’t always see how this affects them. It doesn’t affect funding, it doesn’t affect the amount of allied health professionals that you have, it has no concrete affect on their practice, other than whether their patients are happy or not. (ID = 015)

- And we don’t have to get two doctors to agree on anything to actually make it happen because it’s the community-based nature of this. So decisions are really made taking into account how it affects not just the doctors, but the ITs and everyone else. Really the decisions about going along with D2D were really mostly determined by the direction that Name-X thought we should take and then passed down rather than the other way around. (ID = 014)

- huge changes would be more difficult just because we’re not really allowed to tell them how to work. Basically we have to try it with one physician and then say, hey, you know what, this worked really well, look at the difference in his numbers from doing this for a couple of months. And even so, it’s the same physician every time, so I think some of the doctors get a little, well, I don’t want to hear that from him anymore. We actually do have three physicians on our quality committee now which is great because I think they’ll be willing to try more things too and spread them to the physicians they work with. So we have multiple sites of physicians so that makes it difficult too. (ID = 006)

- We have a number of physicians who are closer to retirement or slowing down or getting out of their practices. They just don’t really have the enthusiasm to implement changes or to try to do something in a different way, whether it’s changing the way that they report something in the EMR to like I’ve been doing it by paper for 30 years of my life. And then I finally converted to the EMR eight years ago and you’re not going to tell me how to do something different for two years before I retire. And then the other factor is just time. As I mentioned earlier, many of our physicians they work in the Emergency Department for a smaller community hospital. So they work in the ED, they’re working on the floors, they’re seeing patients, they’re working in long-term care and they just don’t have the time. They have their own admin time to work on or they have their own clinic time where they have to be here doing that, so that’s kind of one of the struggles. And some individuals see the family health team as being the family health team and the physician group being the physician group and we work together, but we don’t have to play together kind of thing. (ID = 019)
<table>
<thead>
<tr>
<th>Theme</th>
<th>CFIR Domain: Intervention Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of Audit and Feedback Cycles</td>
<td>The time between the release of D2D 2.0 and audit of D2D 3.0 was too short. FHTs did not have enough time to develop any quality improvement initiatives.</td>
</tr>
<tr>
<td>Data: Timing of Measures</td>
<td>Understanding of the timing of the respective data within D2D was inconsistent. The default assumption was that the data were recent as of the release of the report.</td>
</tr>
<tr>
<td>Data: Sensitivity</td>
<td>Many measures were representative of physician behaviour, not FHT behaviour. This lowered the utility of the feedback to FHTs.</td>
</tr>
<tr>
<td>Data: Duplication</td>
<td>In most cases, duplication of data was cited as a barrier to implementation as it provided no additional utility to FHTs as it was not considered to be up-to-date.</td>
</tr>
<tr>
<td>Data: Definitions</td>
<td>Technical definitions of D2D indicators were not clear or easily discoverable. As a result, FHTs questioned the legitimacy of the results.</td>
</tr>
<tr>
<td>Data: Peer Comparison</td>
<td>While appealing as a feature, FHTs raised concerns regarding the validity of the peer comparison in D2D. Little support documentation was available to clarify this.</td>
</tr>
<tr>
<td>Website Design</td>
<td>Three issues were identified with respect to website usability: i) The website was difficult to navigate. ii) The presentation of data was not consistent with conventions in data visualization. iii) Sharing the feedback report or exporting data was not possible, leading to alternate efforts to disseminate the feedback.</td>
</tr>
</tbody>
</table>

**CFIR Domain: Outer Setting**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIDSS Dependence</td>
<td>QIDSS were cited as the implementation leader for D2D. Leading all activities related to participation. In the event a QIDSS was not available, participation would cease or be limited to the minimum.</td>
</tr>
</tbody>
</table>

**CFIR Domain: Inner Setting**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Priority</td>
<td>Relative to the other performance feedback reports available, D2D did not provide enough value to FHTs. As a result, resources were not shifted to D2D from the other activities.</td>
</tr>
<tr>
<td>Available Resources: Time</td>
<td>Participation in D2D, particularly the audit stage, was a significant time burden. As QIDSS are implementation leaders, they lacked the available time to lead much beyond data submission.</td>
</tr>
<tr>
<td>Available Resources: Human Resources</td>
<td>Some FHTs lacked the staffing to support the complete implementation of D2D.</td>
</tr>
<tr>
<td>FHT-FHO Relationship</td>
<td>Physicians saw the FHT as an entity to which they were not accountable.</td>
</tr>
</tbody>
</table>
Discussion

This thesis explored the extent to which team-based primary care practices implemented the data-2-decisions audit and feedback program and the factors which supported or hindered that process. At this stage in its development, D2D appeared to be representative of an incomplete feedback loop, meaning it did not represent the iterative, self-regulating process anticipated when the intervention is implemented successfully. This incomplete feedback loop may have been a product of the intervention itself, the FHTs’ dependence on the QIDSS resource to facilitate implementation, and/or the limited ability of this external resource to accomplish this goal. Further complicating matters was the relationship between the Family Health Team and the physicians it is intended to support. The extent to which each of these main findings contributed to implementation in the context of the broader literature may provide some clarity in assessing their meaning.

Intervention Characteristics

Results from the interviews revealed several limitations to the D2D intervention which restricted its utility to Family Health Teams. This is despite the fact that D2D adheres to many relevant recommendations meant to support the optimization of audit and feedback interventions (64). For example, through the use of peer comparison, D2D satisfied the recommendation regarding the use of comparators to reinforce desired behaviour change (64). Second, D2D satisfied the recommendation of providing multiple instances of feedback to participants – with D2D 3.0 representing the third feedback report during a period of fifteen-months (24,64). In the context of other Audit and Feedback interventions in primary care, D2D is in the minority by providing multiple iterations of feedback (31–43).
Little is known about the impact of the frequency of multiple feedback cycles on implementation. For example, the study by Hysong et al., identified an association between timely feedback (defined as monthly or more frequent) and higher performance (29). However, this study evaluated internal audit and feedback programs at the facilities recruited and not a standardized external audit and feedback program like D2D (29). As a result, it is unclear whether the differences in the high and low performing practices are attributable to the respective internal audit and feedback programs or if some other issues may be at play (29).

In the present study, a high frequency of feedback reports (as perceived by participants) was viewed as a barrier to implementation as it failed to consider the time commitment necessary to conduct the audit or respond to feedback. It is plausible that the extent to which the audit process is automated may influence a practice’s acceptance of a high frequency of feedback reports. However, even if audits are automated, developers of feedback reports must consider the broader impact of higher frequency of feedback as an increasing frequency of feedback may not necessarily yield greater results. For example, previous studies in primary care evaluating A&F interventions (Table 1) demonstrated limit effectiveness despite a range of frequencies of feedback iterations ranging from one-time to quarterly (31–43). In the case of audit and feedback reports with manual or automated auditing, the frequency of feedback reports may be vulnerable to the law of diminishing returns. Within a given time-frame, as the frequency of feedback increases, the marginal benefit of each additional feedback report will fall. Manual auditing may be more sensitive to this effect given the time required to collect the necessary data. This may have been the case with D2D, where it could be argued that feedback fatigue was beginning to emerge among FHTs interviewed for this study. This fatigue was observable when informants expressed frustration that there was not enough time to respond to the feedback and were
subsequently relieved with the delay of the upcoming D2D 4.0 report. Further research is required to understand the impact of the frequency of feedback cycles on implementation and the potential modifying effects of an automated audit process.

In addition to the relationship between automation and feedback fatigue, the recommendation that feedback be provided over multiple iterations also includes an inherent assumption that data are representative of current performance. This is consistent with the recommendation that feedback be provided to its intended audience as soon as possible (64). However, as demonstrated in Figure 2, D2D failed to adhere to this recommendation due to delays inherent in some aspects of the multi-sourced nature of the data used to construct the D2D feedback report. This was most pronounced with D2D 3.0, where ICES data were duplicated from D2D 2.0, as more recent data were not available. This is an important issue in the design of the audit and feedback interventions and current data are necessary to fulfill the iterative, self-regulating nature anticipated of a successful A&F program. However, failure to adhere to this recommendation was an intentional decision in the development of D2D by AFHTO, as it was thought that it would minimize the extraneous cognitive load for feedback recipients. Results from this study suggest that this decision produced barriers which were not previously anticipated (64). For example, the feedback report lacked specific information regarding the relative timing of different measures (Figure 3) meaning that users were not informed that data were not representative of current performance. This, in turn, led feedback recipients to question the validity of the feedback when they compared the data to their own internal data sources.

Efforts to minimize the extraneous cognitive load are cited as another important recommendation to necessary to facilitate the optimization of audit and feedback interventions. In addition to the case above, a 2016 paper by Brown et al. emphasize the important relationship
between good design of a feedback report and its usability for the purposes of meeting its stated objective (65). Results from the usability testing for this thesis revealed that extraneous cognitive load was not minimized. Rather, while the actual web-page had an appealing visual design, its functional design may have interfered with the implementation of the intervention. For example, as noted above, while D2D presents data from multiple sources representing different time periods, none of this information is presented clearly on the feedback report. Further, the absence of a link to the data dictionary on the feedback report served as a barrier to implementation as users were not particularly interested in navigating through the support pages to find the information they needed. However, others were keen on using this information in an attempt to validate the data as they were skeptical of its difference from their other reporting.

Efforts to validate the data in an external audit and feedback report may be a function of the intervention, and may be part of initial efforts to facilitate normalization of an early stage audit and feedback program. Specifically, this may be an element of a setting’s effort to establish coherence with an external audit and feedback program – as posited by Normalization Process Theory (NPT) (66). As described by NPT, coherence is the “sense-making” work that people do when attempting to operationalize a new set of practices (66). As a result, it should be expected that feedback recipients will attempt to validate (or question the validity) of the data from an external feedback report any time the data indicate that changes to processes may be needed. Therefore, the finding from this investigation that FHTs questioned the validity of the data in D2D should be expected and is not a unique feature of D2D. For example, in an evaluation of a different A&F program in the FHT setting Ivers et al. found that physicians also expressed skepticism with respect to the validity of the data presented (30).
The recommendation to optimize an audit and feedback intervention by limiting the cognitive load on feedback recipients may be part of an effort to facilitate the coherence process as described by NPT (64,66). Efforts should be made to provide feedback recipients with the information they need to interpret and potentially respond to feedback and establish coherence with the external feedback program. This may provide an additional explanation as to why A&F has had limited success in the primary care context, as the majority of studies reviewed included external audit and feedback programs. Thus, in addition to the short time frames in which the studies were conducted, recipients may have been hesitant to accept the results as they had yet to establish “coherence” with the audit and feedback cycle (30,33,35–40,43,67). This may also explain why the PPRNet study remains the exception as coherence had already been established between feedback recipients and the external audit and feedback program in which they were actively participating (31,32). Ideally, as feedback recipients establish coherence with an external A&F program, efforts will be reoriented from validation to an assessment of whether or not a response to the feedback is necessary – thus completing the feedback loop (66).

Several opportunities remain for D2D to incorporate further recommendations based on evidence from the literature to optimize feedback design and delivery. For example, data in D2D were not linked with any summary message (64). Participants were required to interpret their data independently based on their interpretations of the data visualizations or their supplemental reports prepared by a FHT’s respective QIDSS. In addition, D2D relied on a single-mode of feedback delivery – even though it is recommended that feedback delivery should be multi-modal (64). Exploration of multi-modal feedback delivery may enhance receptivity of the report and expand the primary user group beyond the QIDSS as currently identified.
Clarifying the Objective of Audit and Feedback

Results from this study suggest that it is not clear what the primary objective of the D2D audit and feedback program is supposed to be. For example, a commonly cited reason why FHTs participated in D2D was to help AFHTO advise the Ministry of Health and Long Term Care (MoHLTC) on two different policy fronts. First, participants saw D2D as a mechanism by which AFHTO could advise the MoHLTC on the direction of quality improvement policy in primary care. FHTs are concerned that they are being evaluated using measures which may not be sensitive to capturing the relevant effects of team-based practice as they rely on physician billing. As a result, D2D may be an attempt to demonstrate the principles which should be considered regarding quality improvement policy in primary care. These are the same principles that inform the potential promise of D2D: measures that are sensitive to FHT practice; informed by participant feedback; comprised of recent data; and meaningful to both primary care providers and patients. Second, participants saw D2D as a mechanism by which AFHTO could advise the MoHLTC on the value of Family Health Teams and their contributions to the health system at large. This may have influenced the development of the “Quality Roll-Up” measure mentioned in the methods section but omitted from the analyses of this thesis. As D2D also includes some cost data, attempts may be made to justify Family Health Teams by comparing “Total Quality” to Cost. However, there are serious limitations to this approach given the “Quality Roll-Up” measure has yet to be validated externally.

With respect to the inner setting, it was surprising to see that only a small number of teams interviewed were participating in D2D to support their efforts with respect to quality improvement. While there were many teams who participated in D2D not interviewed for this study, the observation that so few intended to use it to support QI was surprising. The finding
that a larger proportion of FHTs participated in D2D to support AFHTO’s policy advisory efforts, rather than internal QI efforts, may be one additional explanation as to why the implementation of D2D was representative of an incomplete feedback loop.

Upon examination of the AFHTO website, including the relevant support documentation, no specific objective statement could be found for D2D. Interestingly, in examining the 2016 paper citing recommendations for improving feedback, the inclusion of an explicit objective statement was also absent (64). However, the inclusion of an explicit objective statement may provide greater clarity for the feedback recipient regarding their participation in the intervention and the perceived benefits. For example, the PPRNet performance feedback report has an explicit objective to “improve the delivery of primary care services in a broad spectrum of conditions commonly encountered” (31,32). The objective is clear in that its purpose is to support participants’ efforts regarding quality improvement (31,32). A prospective practice deciding to participate in the A&F program could also judge very quickly whether or not the objective of the intervention is in line with the organizational objectives of the practice.

Further research is necessary to determine the extent to which the presence of an explicit objective for an audit and feedback intervention would facilitate participation and implementation. In addition, while D2D lacks a specific objective, it may still be possible to achieve multiple simultaneous objectives within the Outer and Inner Settings, respectively. However, this may be challenging unless the promise of D2D is fulfilled, as policy advisory efforts will be undermined from a lack of utilization by D2D participants.

**Relative Priority**

The relative priority of D2D was attributable to factors relating to the intervention characteristics, and also the associated incentives with participating in the intervention. As data-
mature FHTs did not identify a relative advantage of D2D over their internal reporting, there was little incentive to invest resources in D2D. Meanwhile, the data immature FHTs lacked the resources to support any potential response to a performance gap.

The trade-offs made by FHTs may provide further clarity as to why D2D did not represent a complete feedback loop. Specifically, FHTs must consider the benefit they receive in return for the resources necessary to participate. With respect to the Outer Setting, support from an external resource in the form of the Quality Improvement Decision Support Specialist (QIDSS) was necessary to participate in D2D. However, the impact of the QIDSS on the Inner Setting was mostly limited to the Audit and Feedback stage of the feedback loop and not the response stage. The inability of the QIDSS resource to support the response stage of the audit cycle may be attributable to the nature of the position, where the individual supports a group of FHTs on a part time basis. Given this, one must question what value the QIDSS could be contributing to a FHT to support D2D if they didn’t have to spend such a considerable amount of time supporting D2D.

The availability of this external resource to support implementation is a feature which is unique to the FHT setting and D2D, as none of the other studies reviewed mentioned a similar resource. However, in many of the studies reviewed, a member of the research team often conducted site visits to review the feedback with the participants and support the response stage (30,33,35–40,43,67). The QIDSS position may be one way to formalize this approach without having to rely on participation in a research project. Nevertheless, the time intensive nature of the audit stage of the feedback loop may have prevented the QIDSS from supporting a potential response stage. Given that D2D lacks an automated auditing process, FHTs are unable to reallocate resources (QIDSS) from the audit stage to the response stage.
This may provide another explanation why the PPRNet study observed greater success compared to the other primary care studies reviewed. PPRNet employed the use of an automated auditing process to support the delivery of feedback reports on a quarterly basis (31,32). In comparison, Family Health Teams were required to conduct manual audits which typically included data abstraction from relevant feedback reports as well as the validation, refinement and potential troubleshooting of EMR queries. As a result, one could assume that the staff at the PPRNet practices could allocate their resources more efficiently by supporting the response stage of the audit cycle.

While the resource trade-off to participate in D2D is low compared to PPRNet, it may be representative of a typical attempt at audit and feedback in primary care. Of the literature reviewed for this investigation, eight of the twelve studies conducted their audit stage manually (31–43). Of the four studies which had automated EMR abstraction processes (Table 1), only two represented A&F programs which were implemented in a natural practice context (31,32,42). Further, the studies with manual data abstraction may have faced similar challenges to implementation from the perspective of the inner setting as were observed with the implementation of D2D (33,35–40,43,67). Therefore, the resources required to complete the audit stage may interfere with a FHT’s ability to implement subsequent stages of the feedback cycle. In the long run, this may serve as a barrier to normalization – unless investments are made to improve the technical efficiency of the audit stage of an audit and feedback program.

**FHT-FHO Relationship**

At the time of this writing, few studies have explored the impact of the relationship between provider groups in team-based primary care on the implementation of a quality improvement initiative. However, there may have been emergent signs in previous work of this
tension. For example, in the 2014 study by Ivers et al., physicians in family health teams cited concerns about balancing their time between patient centred care and quality improvement. In addition, participants openly questioned whether it was an appropriate use of physician time to engage in quality improvement (30). After revisiting the literature, no studies have explored this relationship with respect to audit and feedback in particular. The finding from this study that the FHT-FHO relationship may serve as a barrier and a facilitator to implementation opens to the door to a previously unexplored research gap.

Data from this investigation may provide some early direction in identifying factors which may explain when and in what contexts the FHT-FHO relationship will be a barrier to the implementation of an audit and feedback program. Of the FHTs interviewed, the FHT-FHO relationship was cited as both a barrier and as a facilitator to quality improvement – with the former reported more frequently. In situations where the FHT-FHO relationship was cited as a barrier to implementation, it was specifically noted that the FHO, or even individual physicians, did not see the FHT as an entity to which they were accountable.

One must also question if the model itself may produce an environment supportive of the findings of this investigation. The FHT model is unique in the way it protects physician practice autonomy through the establishment of a partnership between the FHT and the FHO, rather than having the physicians as FHT employees or contractors. Considerable effort may be required on the part of the FHT to engage with the FHO as a whole, or even individual physicians. This is perhaps best demonstrated by the observation where physicians’ lack of interest in D2D was a function of incentive. It was suggested that alternative A&F initiatives, such as the CCO-SAR were more utilized as physicians could use it to maximize their income – although these claims were not validated for this study. Physician engagement in quality improvement may also be a
function of time. For example, older more established FHTs may have had more time to encourage attitude and behaviour change to support quality improvement.

Further research is necessary to understand the relationship between team-based primary care models and physician willingness to support quality improvement in team-based primary care settings. Of specific interest would be two different methodological approaches examining the physician-team relationship to support audit and feedback. First, a within model study could explore the findings of this investigation in greater detail. Validation of the observation that the FHT-FHO relationship may serve as a barrier to implementation of audit and feedback is necessary with a much larger sample than that used in the present study. In addition, this hypothetical study could identify specific FHT characteristics and structures in which the magnitude of the observed barrier is greatest. Second, a between model study would be necessary to determine if the physician-team relationship’s impact on the implementation of audit and feedback is exclusive to Family Health Teams or other models of team-based primary care. Of specific interest would be a comparison between the Family Health Team and the Primary Care Network in Alberta given that the standard tri-lateral governance model of the PCN includes representation from the public insurer (the payer) as well as the local regional health authority (Alberta Health Services) (68). This is in contrast from Family Health Teams which are typically governed by physician or community boards and have no representation from the public insurer or local regional health authority. Alternatively, one could explore this relationship within the Patient Centred Medical Home model in the United States. This between models approach could test the hypothesis that physician support of team-based quality improvement activities is a function of the governance model (including representation from other stakeholders) of the team-based primary care practice. Understanding the mechanisms by
which physicians can work with other interdisciplinary health providers in a team-based primary care environment to support Quality Improvement is necessary to understand how to truly optimize the effectiveness of an intervention such as audit and feedback.

**Deployment of CFIR Methodology**

The findings from the present study were elucidated through the use of the Consolidated Framework for Implementation Research (CFIR) (55). It provided a pragmatic structure which was used to guide data collection and analysis. This was perhaps best demonstrated through the exclusion of the domain “Characteristics of Individuals” and its associated constructs. As this construct were not considered to be feasible to be included within the parameters of this thesis, it was excluded as enabled by the guideline. Most importantly, the exclusion of domains or constructs from the investigation did not affect the validity of the findings nor the interpretations of research. The domains which were included in the study (Intervention Characteristics, Inner and Outer setting, and Process) enabled the collection of rich data from participants regarding the implementation of the audit and feedback intervention. This study should continue to add to the growing literature surrounding the use of the CIFR to evaluate the implementation of a variety of health interventions in a range of contexts.

**Limitations**

There are a number of limitations to this study which must be considered. First, participants were relatively early adopters of D2D and many had roles on the AFHTO board or the development of D2D. If these FHTs were replaced with others which had no AFHTO association beyond membership, some of the emergent findings may have been different – particularly surrounding the decision to participate in the intervention. Second, participants were recruited from a homogenous sample limiting the generalizability of the findings to team-based
primary care settings. Further, given the unique nature and structure of the Family Health Team practice model – generalizability to other team-based practice settings may also be limited. As described above, further research is required to determine if the observed relationship between the FHT and the FHO is unique to the Family Health Team model or not. Third, while analysis of transcripts was supported by double coding of interviews, thematic coding was done independently with no validation. Given the explicit statement of application of the framework methodology in concert with the use of deductive coding, the impact of this limitation on the findings from this investigation should be marginal. Further, the use of a previously validated framework should support the prospective reproducibility of this investigation.

**Conclusion**

Findings from this study suggest that the implementation of the data-2-decisions audit and feedback program is representative of an incomplete feedback loop. Several barriers were identified which prevented Family Health Teams from closing the feedback loop and completing implementation. First, the nature of the audit process was time intensive, forcing FHTs to allocate their resources to data abstraction. Second, in most cases implementation was led by an external resource supporting multiple FHTs simultaneously on a part time basis. This limited their ability to promote engagement. Third, FHTs had little incentive to execute the intervention citing data validity concerns in concert with superior internal reporting practices. Fourth, the relationship between the FHT and independent physician group was cited as both a barrier and facilitator to implementation. Given the independence of the physician group, the FHT had few tools at their disposal to entice participation in quality improvement efforts. Despite the many barriers to implementation, key informants cited the promise of a superior audit and feedback tool representative of user feedback and current performance as a key facilitator in deciding to
participate. Further, informants also cited their participation as their support of efforts to use the A&F tool to advise policy on quality improvement in primary care as well as the value of the Family Health Team model. The presence of these facilitators to participation in concert with the absence of an explicit objective for the A&F program may delay the fulfillment of the promise of the intervention. With an explicit objective, the barriers to implementation may be addressed enabling the completion of the feedback loop by aligning the priorities of the developers and users of the intervention. Further research should investigate how audit and feedback interventions are implemented in natural practice settings to identify the conditions where the intervention works best. The findings from the present study may serve as a launching point for these efforts in team-based primary care.
References


46. QSR International PTy Ltd. NVivo qualitative data analysis Software.


Appendices

Appendix 1. Ethics Documentation

PROTOCOL REFERENCE # 31773

February 4, 2016

Dr. Frank Sullivan
DEPT OF FAMILY & COMMUNITY MEDICINE
FACULTY OF MEDICINE

Dear Dr. Sullivan,

Re: Your research protocol entitled, "Developmental evaluation of the past and future potential impact of a performance measurement initiative on beliefs, behaviours and outcomes related to Quality Improvement (QI) and Performance Measurement (PM) in primary care"

We are writing to advise you that a member of the Health Sciences Research Ethics Board (REB) has granted approval to an amendment (Received January 13, 2016) to the above-referenced research protocol under the REB’s delegated review process. This amendment approval letter only applies to what was outlined in the request form under section 5.a) or otherwise marked in the revised protocol.

Any changes to the approved protocol or consent materials must be reviewed and approved through the amendment process prior to its implementation. Any adverse or unanticipated events should be reported to the Office of Research Ethics as soon as possible.

Best wishes for the successful completion of your research.

Yours sincerely,

Elizabeth Peter, Ph.D.
REB Chair
PROTOCOL REFERENCE # 32242

February 4, 2016

Dr. Frank Sullivan
DEPT OF FAMILY & COMMUNITY MEDICINE
FACULTY OF MEDICINE

Dear Dr. Sullivan,

Re: Your research protocol entitled, "Establish and estimate impact of thresholds for performance in the calculation of a composite quality indicator for primary care"

We are writing to advise you that a member of the Health Sciences Research Ethics Board (REB) has granted approval to an amendment (Received January 13, 2016) to the above-referenced research protocol under the REB’s delegated review process. This amendment approval letter only applies to what was outlined in the request form under section 5.a) or otherwise marked in the revised protocol.

Any changes to the approved protocol or consent materials must be reviewed and approved through the amendment process prior to its implementation. Any adverse or unanticipated events should be reported to the Office of Research Ethics as soon as possible.

Best wishes for the successful completion of your research.

Yours sincerely,

Elizabeth Peter, Ph.D.
REB Chair
Dear Mr. Wagner,

Thank you for your e-mail. You are not required to submit any documentation for review by the Women’s College Hospital Research Ethics Board. Ethics approval from U of T for the research project described below is sufficient.

Best regards,

Melissa

Dear Melissa,

Please allow me to introduce myself. My name is Daniel Wagner, I am a graduate student at the Institute of Health Policy, Management and Evaluation at the University of Toronto. I currently work under the supervision of Dr. Noah Ivers at Women’s College Hospital. I am writing you today to seek clarification on WCHRI’s ethics requirements for my research project.

The project I am working on is part of a larger research program run out of the University of Toronto and has already received ethics approval from UofT. When Dr. Ivers and I joined the project, we submitted an amendment to add our names to the relevant protocols. I should also mention that the PI on the project is Dr. Frank Sullivan at UofT. I should also mention that the project will not involve any patients from WCH, and data collection will consist of semi-structured interviews with executive directors of family health teams in Ontario.

Given Dr. Ivers’ role at WCH and my position as a graduate trainee at WCHRI, we wanted to clarify the requirements of WCHRI Ethics before we proceed with data collection. Are we required to submit documentation for review?

Any guidance you have on this matter would be greatly appreciated.

Regards,

Daniel Wagner
E: dan.wagner@mail.utoronto.ca
M: 403-968-3010

CONFIDENTIALITY NOTICE: This e-mail message, including any attachments, is for the sole use of the intended recipient(s) and may contain confidential and privileged information. Any unauthorized review, use, disclosure or distribution is prohibited. If you are not the intended recipient, please contact the sender by reply e-mail and destroy all copies of the original message.
## Appendix 2. D2D Data Submission Form

### DATA TO DECISIONS – DATA INPUT

This data has already been submitted and may not be altered.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost</td>
<td>100.0</td>
</tr>
<tr>
<td>Cost - Primary Care</td>
<td>100.0</td>
</tr>
<tr>
<td>Cost - Services</td>
<td>100.0</td>
</tr>
<tr>
<td>Cost - Settings</td>
<td>100.0</td>
</tr>
<tr>
<td>Cost - Institutions</td>
<td>100.0</td>
</tr>
<tr>
<td>Patients served</td>
<td>100.0</td>
</tr>
<tr>
<td>Patients rostered</td>
<td>100.0</td>
</tr>
<tr>
<td>SAME score</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Numerator</th>
<th>Denominator</th>
<th>Rate (%)</th>
<th>Numerator</th>
<th>Denominator</th>
<th>Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient experience survey</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data from notes (access through eHR or primary care practice input)</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childhood vaccinations</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes care</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EHR Data Quality (expressed as a numerical, e.g., 0.75)</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician Capacity - office appointments</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician Capacity - other activities</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physicians represented in estimate (as % total on team)</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHP Capacity - office appointments</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHP Capacity - other activities</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description of team's approach to Capacity Indicator**
Appendix 3. Interview Recruitment Letters

Dear [insert name of ED]

With the recent release of D2D 3.0, we are curious to learn more about how teams are using the tool.

Given your role in your FHT and your teams’ performance on the quality measures included in D2D – we are particularly interested in your thoughts about the role D2D plays in your team’s ongoing quality improvement activities. This includes whether or not it has been helpful and also what might be done to enhance the impact of the program in the future.

We have partnered with the research community, specifically UPTOPIAN (University of Toronto’s Practice-based Research Network) and Daniel Wagner, a Master’s Student at the Institute of Health Policy, Management and Evaluation at the University of Toronto. This work is being conducted as part of Mr. Wagner’s thesis research project and is therefore independent of AFHTO. For the purposes of this project, AFHTO will only have access to the final report where all interview data will be summarized and anonymized. Mr. Wagner is conducting his research under the supervision of Dr. Noah Ivers – a leading expert in the field of Audit and Feedback.

We want you to feel comfortable sharing your experiences with D2D.

We hope you might be able to find approximately 45 minutes to talk with Mr. Wagner in March and April 2016. While we understand that this is a busy time for you as fiscal-year end approaches, it is our hope to get your feedback closer to the release of D2D 3.0 – when it will be top of mind.

If you are interested in having an interview with Mr. Wagner to discuss your team’s use of D2D, please provide your contact information using the secure link below. Mr. Wagner will then follow-up with you to schedule your interview.

[LINK TO SECURE REGISTRATION FORM]

Thank you in advance for taking the time to provide your feedback.

Angie Heydon, ED
AFHTO
Dear Colleague,

I am writing you today regarding an exciting opportunity to contribute to a project exploring Family Health Teams’ engagement with the Data-2-Decisions program run by AFHTO. As Executive Directors of FHTs, we are intended to be the primary recipients of the feedback within the D2D tool. Your participation will consist of a one-on-one interview with Daniel Wagner, a graduate student at the Institute of Health Policy, Management and Evaluation at the University of Toronto.

This work is being conducted as part of Mr. Wagner’s thesis research project. As a result, while Mr. Wagner does have a partnership with AFHTO, this work will be done independently of the organization. In addition, Mr. Wagner is conducting his research under the supervision of Dr. Noah Ivers – a leading expert in the field of Audit and Feedback, and also one of our family physicians here at the Women’s College Family Practice. In speaking with both Mr. Wagner and Dr. Ivers I have been ensured that all research data belongs to them and not AFHTO. Further, I have been informed that all interview data will be anonymized in any final publication or report.

Each interview will take approximately 45 minutes and will consist of two parts. The first part will be a conversation about your team’s use of the D2D feedback report. The second part will be a short session exploring the usability of the D2D website. With the recent release of D2D 3.0, now is the ideal time to explore how Family Health Teams are actually using the report.

If you are interested in meeting with Mr. Wagner, please provide your contact information using the secure link below. Mr. Wagner will then follow-up with you to schedule an interview.

[LINK TO SECURE REGISTRATION FORM]

Thank you in advance for taking the time to provide your input.

Ken Callaghan, BA, MBA
Executive Director | Academic Family Health Team
Manager | Family Practice Health Centre
Women’s College Hospital
77 Grenville Street, Rm 214 | TORONTO, Ontario M5S 1B2
T 416.323.6400 ext 4595  F 416.323.7323
ken.callaghan@wchospital.ca | www.womenscollegehospital.ca

118
Appendix 4. Interview Registration Form

Register for an Interview

Thank you for your interest in meeting with me to discuss how your Family Health Team uses the feedback provided as part of the Data-Driven Decisions program.

In the fields below, please provide your name and contact information. The information you provide is stored on a secure server and will not be shared with anyone outside this research project. I will contact you within one business day to schedule your interview.

I would like to thank you in advance for taking the time out of your busy schedule to contribute to my research.

Daniel Wagner
MSc Student
Institute of Health Policy, Management and Evaluation
University of Toronto
E: dan.wagner@mail.utoronto.ca

What is your name?

What is the name of your organization?

Are you the Executive Director (or equivalent position) at your Family Health Team?

No

What is your phone number?

XX-XXX-XXXX

What is your email address?

What is your preferred method of contact?

What is your DFD Team code?

This is the identifier which you enter into the DFD report to link your team’s data. Having this information will help me better understand your team’s characteristics.

Is there anything you would like me to be aware of prior to contacting you?

Feel free to share any information you feel is important.

Submit

Never give out your password. Don’t give your personal information to companies you don’t trust.
Appendix 5. Qualitative Interview Guide

<table>
<thead>
<tr>
<th>Interviewee:</th>
<th>Interview Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FHT Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interview Type:</th>
<th>In-Person</th>
<th>Skype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview Type:</td>
<td>In-Person</td>
<td>Skype</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject Role in FHT:</th>
<th>ED</th>
<th>Other:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Role in FHT:</td>
<td>ED</td>
<td>Other:</td>
</tr>
</tbody>
</table>

Introduction

Hello,

Thank you for agreeing to participate in this study and to speak with me today. My name is Daniel Wagner, I am a graduate student at the Institute of Health Policy, Management and Evaluation at the University of Toronto. I am interested in learning about how Family Health Teams engage with data in performance feedback reports, such as the Data-2-Decisions (D2D) program.

This two-part interview should take no more than 60 minutes and will be audio-recorded to ensure that all of the key points you share are captured accurately. The first part will be a conversation about your team’s use of the D2D feedback report. The second part will be a short session exploring the usability of the D2D interactive web site.

The results of our conversation will be incorporated into my final thesis project, however your comments will be de-identified. While AFHTO will receive a copy of the final report, it will not include any identifiable information. You can be assured that your comments during our conversation are confidential.

Before we begin, there are three items I would like to review:

First, do you have any questions about this research?

Second, I would like you to know that any identifying information that may be disclosed during this discussion (such as team, health professional, or patient names), will be removed from interview transcripts.

Third, at any point during the interview you are free to end the discussion or withdraw from the study entirely.

In-person Interview: Yes Time Acknowledged: ________________

In-person Interview: No Time Acknowledged: ________________
Background & Rationale

<table>
<thead>
<tr>
<th>Question</th>
<th>Action from Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which iterations of D2D has your FHT participated in?</td>
<td></td>
</tr>
<tr>
<td>- D2D-V1</td>
<td></td>
</tr>
<tr>
<td>- D2D-V2</td>
<td></td>
</tr>
<tr>
<td>- D2D-V3</td>
<td></td>
</tr>
<tr>
<td>Why does your FHT choose to participate in D2D?</td>
<td></td>
</tr>
<tr>
<td>PROBE: Is it helping you?</td>
<td></td>
</tr>
<tr>
<td>PROMPT: Help me understand the value D2D provides to your team.</td>
<td></td>
</tr>
<tr>
<td>Who accesses and reviews the results of the D2D Report?</td>
<td></td>
</tr>
<tr>
<td>PROBE:</td>
<td></td>
</tr>
<tr>
<td>How did you come to this decision?</td>
<td></td>
</tr>
<tr>
<td>Has this changed over time? Why?</td>
<td></td>
</tr>
<tr>
<td>Have you ever viewed the results of the report?</td>
<td></td>
</tr>
<tr>
<td>Can you explain how the results of the D2D report are used within your FHT?</td>
<td></td>
</tr>
<tr>
<td>PROBE:</td>
<td></td>
</tr>
<tr>
<td>Are any actions taken in response to the results? If yes, please describe.</td>
<td></td>
</tr>
<tr>
<td>PROMPTS:</td>
<td></td>
</tr>
<tr>
<td>Is it shared with other staff?</td>
<td></td>
</tr>
<tr>
<td>How are the results reviewed? With who?</td>
<td></td>
</tr>
<tr>
<td>Who else is involved with leading your FHT’s participation in D2D?</td>
<td></td>
</tr>
<tr>
<td>How did you come to this decision?</td>
<td></td>
</tr>
<tr>
<td>Are you aware of other organizations participating in D2D?</td>
<td></td>
</tr>
<tr>
<td>Does this impact the support for D2D in your setting? If so, how?</td>
<td></td>
</tr>
</tbody>
</table>

Value

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you perceive a strong need for D2D? Why or why not?</td>
</tr>
<tr>
<td>Are your thoughts consistent with other members of your team?</td>
</tr>
<tr>
<td>If not, how are they different? Does that impact the utility of D2D?</td>
</tr>
<tr>
<td>Besides D2D, does your FHT engage in any other performance feedback activities? [If No:] Why not? Do you believe that participation in D2D is sufficient? [If Yes:] How does D2D compare to the other interventions you mentioned? What advantages does D2D have compared to the other performance feedback program? What disadvantages does D2D have compared to other performance feedback programs? Are they complimentary? Do you rely on the information from one more than another?</td>
</tr>
<tr>
<td>Is there another tool that your FHT would prefer to use instead of D2D?</td>
</tr>
<tr>
<td>Can you describe that program? Why would that program be preferential?</td>
</tr>
</tbody>
</table>
Do you feel that outcome comparisons with other FHTs provide an advantage for your organization?
**PROBE:** If yes, to what extent do those comparisons provide an advantage?
**PROMPT:** What type of advantage does it provide? (e.g., competitive advantage)

What impact has D2D had on your team’s ability to meet, or make progress toward, your organizational objectives?
**PROMPTS:** What specific objectives are now better served?
- What strategies were attempted to meet these goals before D2D?
- What characteristics about D2D enabled your team to overcome this gap?

### Contextual Factors

**Did your FHT consider any evidence or proof when deliberating your participation in D2D?**
**PROBE:** What type of evidence did your FHT consider?
- How did this knowledge affect your participation?
**PROMPT:** Practice guidelines? Published literature? Communication with a colleague?
- Information from thought leaders?

To what extent did the current policy climate for primary care influence your decision to participate in D2D?
**PROBE:** Why?
**PROMPT:** What policy changes do you anticipate, if any?
- How will D2D affect your FHT’s ability to adhere to any upcoming regulatory changes?

How has the structure of your FHT affected your team’s ability to act on the feedback?
**PROMPT:** Has the structure of your team impacted individuals’ engagement with D2D?
- Will your FHT be able to work around these structural challenges?
- Have you needed to do anything to accommodate D2D?
- Will you need to make any additional accommodations moving forward?

### Resources

**How well does D2D participation fit with existing work processes and practices in your FHT?**
**PROBE:** Have any issues or complications arose that were problematic?

Does your FHT have sufficient resources to implement or administer D2D?
**[If Yes]:**
**PROBE:** Is there a minimal amount of required resource? Are there any resources that you received or would have liked to receive?
PROMPT: Could you tell me about what those resources are? What resources have been easy to acquire?

[If No]:

What resources are available to you? What resources do you think you would require to support the implementation/uptake of D2D?

What kinds of information and materials about the D2D have been made available to you?

Who do you ask if you have questions about D2D? Or loading the feedback report?

Has your organization received assistance from someone (person or organization) to support the implementation and use of D2D?

Can you describe this person/group? What is their role? What kind of activities do they lead? How helpful do you think they are? In what ways?

What kinds of changes do you think need to be made to D2D ITSELF to improve implementation in your practice? Are there any components that should not be altered?

What kinds of changes do you think need to be made to supports that exist alongside D2D to improve implementation in your practice? Are there any components that should not be altered?

[End Part One]

Check time.

We've now completed the first part of the one-on-one interview. I'd like to move on to the next part of our conversation. Do you have another 25-30 minutes?

[If No:] Would you like to schedule another time in the coming days to conduct a usability test of the D2D tool? Your continued participation would be extremely valuable. We also have more flexibility in how we do this, and we could schedule the session to be in-person or remote.

Follow-up Session Choice: In-Person Remote

Meeting Details Date & Time

[If Yes:] Begin Usability Test

[If Skype Interview]: Initiate Screen Sharing

To support this part of the interview, I’ll need you to share your screen with me so that I can observe what you’re looking at while conducting the usability test.

Please click the “+” icon in Skype. You should see a list of options. Select SHARE SCREENS.
I am now able to see what is presented to you on your screen.

[Load D2D Interactive Report]

**Usability Test**

We’re now going to do a few usability exercises with the D2D feedback report to explore the impact of design on use. This should take about 30 minutes.

I would like to make clear that I’m testing the site, not you. You can’t do anything wrong here. As you use the site, I’m going to ask you as much as possible to think out loud. That is, to say what you’re looking at, what you’re trying to do and what you’re thinking.

I would like to stress that I am looking for your honest opinion.

If you have any questions as we go along, just ask them. I may not be able to answer them right away, since we’re interested in how people do when they don’t have someone sitting next to them. I will be happy to address any unanswered questions at the end.

Now I’m going to ask you to complete two specific tasks. I’m going to read each one out loud and give you a printed copy. Again, as much as possible, it will help if you can try to think out loud as you go along.

[Set user browser to D2D interactive Report Page]

**1.1: Load FHT-Specific Results for D2D-V3**

Scenario: You have just received notice that D2D-V3 results have been released. Load the results for your FHT.

Notes: Feel free to interact with the page as you see fit, but do not leave the page. That means do not click on a link, or hit the back/forward buttons in your browser.

Prompt List:

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What would you do next?</td>
<td>How do you think it would work?</td>
</tr>
<tr>
<td>What are you thinking?</td>
<td>What are you trying to do now?</td>
</tr>
<tr>
<td>What are you looking at?</td>
<td>Was there something in particular that made you choose that?</td>
</tr>
<tr>
<td>What are you doing now?</td>
<td></td>
</tr>
<tr>
<td>Is that what you expected to happen?</td>
<td></td>
</tr>
</tbody>
</table>

**1.2: Interpret Results of Indicators**

Scenario: D2D-V3 results have been successfully loaded. You can now view the results. Please review the **CORE D2D measures**.

Prompt List:
Why do you say that? | What do you think this indicator represents?
--- | ---
What do you think that means? | Would you say the results are easy to interpret?
What is the scale presented in the graph? | Where would you look for clarity about the data?
What were you expecting to see? | How would you expect this to work?
Where is your team performing well? | 
What are some areas your team could be performing better? | 
What is your impression of your team’s performance? | 

Wrap-Up

That completes all the questions I have for you today. Is there anything else you would like to add before we conclude the interview?

Thank you very much for taking the time out of your schedule to speak with me. Your input in this project is greatly appreciated.

[End Interview]

For Administrative Use Only:

<table>
<thead>
<tr>
<th>Final Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview Completed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional Notes:</th>
</tr>
</thead>
</table>
## Appendix 6. Qualitative Codebook

Note: This template provides inclusion and exclusion criteria for most constructs. Please post additional inclusion and exclusion criteria, guidance, or questions to the [CFIR Wiki](#) discussion tab in order to help improve the CFIR.

Related Interview questions have been flagged in the Interview Guide. Please use that document for reference.

### Innovation Characteristics

<table>
<thead>
<tr>
<th>A. Evidence Strength &amp; Quality</th>
<th><strong>Definition:</strong> Stakeholders’ perceptions of the quality and validity of evidence supporting the belief that the innovation will have desired outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Inclusion Criteria:</strong> Include statements regarding awareness of evidence and the strength and quality of evidence, as well as the absence of evidence or a desire for different types of evidence, such as pilot results instead of evidence from the literature.</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion Criteria:</strong> Exclude statements regarding the receipt of evidence as an engagement strategy.</td>
</tr>
<tr>
<td>B. Relative Advantage</td>
<td><strong>Definition:</strong> Stakeholders’ perception of the advantage of implementing the innovation versus an alternative solution.</td>
</tr>
<tr>
<td></td>
<td><strong>Inclusion Criteria:</strong> Include statements that demonstrate the innovation is better (or worse) than existing programs. (<em>Also include comments about duplication</em>).</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion Criteria:</strong> Exclude statements that demonstrate a strong need for the innovation and/or that the current situation is untenable and code to Tension for Change.</td>
</tr>
<tr>
<td>C. Adaptability</td>
<td><strong>Definition:</strong> The degree to which an innovation can be adapted, tailored, refined, or reinvented to meet local needs.</td>
</tr>
<tr>
<td></td>
<td><strong>Inclusion Criteria:</strong> Include statements regarding the (in)ability to adapt the innovation to their context, e.g., complaints about the rigidity of the protocol. Suggestions for improvement can be captured in this code but should not be included in the rating process, unless it is clear that the participant feels the change is needed but that the program cannot be adapted. However, it may be possible to infer that a large number of suggestions for improvement demonstrates lack of compatibility, see exclusion criteria below.</td>
</tr>
<tr>
<td></td>
<td><strong>Exclusion Criteria:</strong> Exclude or double code statements that the innovation did or did not need to be adapted to Compatibility.</td>
</tr>
</tbody>
</table>

### Outer Setting

| A. Peer Pressure | **Definition:** Mimetic or competitive pressure to implement an innovation, typically because most or other key peer or competing organizations have already implemented or are in a bid for a competitive edge. |
Inclusion Criteria: Include statements about perceived pressure or motivation from other entities or organizations in the local geographic area or system to implement the innovation.

Exclusion Criteria:

| B. External Policy & Incentives | Definition: A broad construct that includes external strategies to spread innovations including policy and regulations (governmental or other central entity), external mandates, recommendations and guidelines, pay-for-performance, collaboratives, and public or benchmark reporting. |

Inclusion Criteria: Include descriptions of external performance measures from the system.

Exclusion Criteria:

<table>
<thead>
<tr>
<th>Inner Setting</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Structural Characteristics</td>
<td>Definition: The social architecture, age, maturity, and size of an organization.</td>
</tr>
</tbody>
</table>

Inclusion Criteria:

Exclusion Criteria:

| B. Implementation Climate | Definition: The absorptive capacity for change, shared receptivity of involved individuals to an innovation, and the extent to which use of that innovation will be rewarded, supported, and expected within their organization. |

Inclusion Criteria: Include statements regarding the general level of receptivity to implementing the innovation.

Exclusion Criteria: Exclude statements regarding the general level of receptivity that are captured in the sub-codes.

| 1. Tension for Change | Definition: The degree to which stakeholders perceive the current situation as intolerable or needing change. |

Inclusion Criteria: Include statements that (do not) demonstrate a strong need for the innovation and/or that the current situation is untenable, e.g., statements that the innovation is absolutely necessary or that the innovation is redundant with other programs. Note: If a participant states that the innovation is redundant with a preferred existing program, (double) code lack of Relative Advantage, see exclusion criteria below. Exclusion Criteria: Exclude statements regarding specific needs of individuals that demonstrate a need for the innovation, but do not necessarily represent a strong need or an untenable status quo. |
<table>
<thead>
<tr>
<th></th>
<th><strong>Exclusion Criteria</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exclude statements that demonstrate the innovation is better (or worse) than existing programs and code to Relative Advantage.</td>
</tr>
</tbody>
</table>
| 2. **Compatibility** | **Definition:** The degree of tangible fit between meaning and values attached to the innovation by involved individuals, how those align with individuals’ own norms, values, and perceived risks and needs, and how the innovation fits with existing workflows and systems.  
**Inclusion Criteria:** Include statements that demonstrate the level of compatibility the innovation has with organizational values and work processes. Include statements that the innovation did or did not need to be adapted as evidence of compatibility or lack of compatibility. *(Include references to meaningful data, meaningful measures, etc.)*  
**Exclusion Criteria:** Exclude or double code statements regarding the priority of the innovation based on compatibility with organizational values to Relative Priority, e.g., if an innovation is not prioritized because it is not compatible with organizational values. |
| 3. **Relative Priority** | **Definition:** Individuals’ shared perception of the importance of the implementation within the organization.  
**Inclusion Criteria:** Include statements that reflect the relative priority of the innovation, e.g., statements related to change fatigue in the organization due to implementation of many other programs.  
**Exclusion Criteria:** Exclude or double code statements regarding the priority of the innovation based on compatibility with organizational values to Compatibility, e.g., if an innovation is not prioritized because it is not compatible with organizational values. |
| 4. **Organizational Incentives & Rewards** | **Definition:** Extrinsic incentives such as goal-sharing, awards, performance reviews, promotions, and raises in salary, and less tangible incentives such as increased stature or respect.  
**Inclusion Criteria:** Include statements related to whether organizational incentive systems are in place to foster (or hinder) implementation, e.g., rewards or disincentives for staff engaging in the innovation.  
**Exclusion Criteria:** |
| C. **Readiness for Implementation** | **Definition:** Tangible and immediate indicators of organizational commitment to its decision to implement an innovation.  
**Inclusion Criteria:** Include statements regarding the general level of readiness for implementation.  
**Exclusion Criteria:** Exclude statements regarding the general level of readiness for implementation that are captured in the sub-codes. |
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 1. Leadership Engagement | **Definition**: Commitment, involvement, and accountability of leaders and managers with the implementation of the innovation. | **Inclusion Criteria**: Include statements regarding the level of engagement of organizational leadership.  
**Exclusion Criteria**: Exclude or double code statements regarding leadership engagement to Engaging: Formally Appointed Internal Implementation Leaders or Champions if an organizational leader is also an implementation leader, e.g., if a director of primary care takes the lead in implementing a new treatment guideline. Note that a key characteristic of this Implementation Leader/Champion is that s/he is also an Organizational Leader. |
| 2. Available Resources | **Definition**: The level of resources organizational dedicated for implementation and on-going operations including physical space and time. | **Inclusion Criteria**: Include statements related to the presence or absence of resources specific to the innovation that is being implemented.  
**Exclusion Criteria**: Exclude statements related to training and education and code to Access to Knowledge & Information.  
In a research study, exclude statements related to resources needed for conducting the research components (e.g., time to complete research tasks, such as IRB applications, consenting patients). |
| 3. Access to Knowledge & Information | **Definition**: Ease of access to digestible information and knowledge about the innovation and how to incorporate it into work tasks. | **Inclusion Criteria**: Include statements related to implementation leaders' and users' access to knowledge and information regarding use of the program, i.e., training on the mechanics of the program.  
*(Include statements commenting on the accessibility/utility of support resources for the intervention).*  
**Exclusion Criteria**: Exclude statements related to engagement strategies and outcomes, e.g., how key stakeholders became engaged with the innovation and what their role is in implementation. |

**Process**

A. Engaging  
**Definition**: Attracting and involving appropriate individuals in the implementation and use of the innovation through a combined strategy of social marketing, education, role modeling, training, and other similar activities.  
**Inclusion Criteria**: Include statements related to engagement strategies and outcomes, i.e., if and how staff and innovation participants became engaged with the innovation and what their role
is in implementation. Note: Although both strategies and outcomes are coded here, the outcome of engagement efforts determines the rating, i.e., if there are repeated attempts to engage staff that are unsuccessful, or if a role is vacant, the construct receives a negative rating. In addition, you may also want to code the "quality" of staff - their capabilities, motivation, and skills, i.e., how good they are at their job, and this data affects the rating as well.

Exclusion Criteria: Exclude statements related to specific sub constructs.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 1. Opinion Leaders | **Definition:** Individuals in an organization that have formal or informal influence on the attitudes and beliefs of their colleagues with respect to implementing the innovation.  
**Inclusion Criteria:** Include statements related to engagement strategies and outcomes, e.g., how the opinion leader become engaged with the innovation and what their role is in implementation. Note: Although both strategies and outcomes are coded here, the outcome of efforts to engage staff determines the rating, i.e., if there are repeated attempts to engage an opinion leader that are unsuccessful, or if the opinion leader leaves the organization and this role is vacant, the construct receives a negative rating. In addition, you may also want to code the “quality” of the opinion leader here – their capabilities, motivation, and skills, i.e. how good they are at their job, and this data affects the rating as well.  
**Include statements about the FHT engaging, or attempting to engage the FHO group. The FHT-FHO dynamic with respect to D2D and performance feedback/quality improvement should be captured here. FHO Group = Opinion Leaders.**  
**Exclusion Criteria:** |
| 2. External Change Agents | **Definition:** Individuals who are affiliated with an outside entity who formally influence or facilitate innovation decisions in a desirable direction.  
**Inclusion Criteria:** Include statements related to engagement strategies and outcomes, e.g., how the external change agent (entities outside the organization that facilitate change) became engaged with the innovation and what their role is in implementation, e.g., how they supported implementation efforts.  
Note: Although both strategies and outcomes are coded here, the outcome of efforts to engage staff determines the rating, i.e., if there are repeated attempts to engage an external change agent that are unsuccessful, or if the external change agent leaves their organization and this role is vacant, the construct receives a negative rating. In addition, you may also want to code the "quality" of the |
external change agent here - their capabilities, motivation, and skills, i.e., how good they are at their job, and this data affects the rating as well. *(Code any statements related to activities of the QIDSS).*

**Exclusion Criteria:** Note: It is important to clearly define what roles are external and internal to the organization. Exclude statements regarding facilitating activities, such as training in the mechanics of the program, and code to Access to Knowledge & Information *if* the change agent is considered internal to the study, e.g., a staff member at the national office. If the study considers this staff member internal to the organization, it should be coded to Access to Knowledge & Information, even though their support may overlap with what would be expected from an External Change Agent.

**B. Executing**

**Definition:** Carrying out or accomplishing the implementation according to plan.

**Inclusion Criteria:** Include statements that demonstrate efforts to implement any component of the D2D audit cycle.

**Exclusion Criteria:**

**C. Reflecting & Evaluating**

**Definition:** Quantitative and qualitative feedback about the progress and quality of implementation accompanied with regular personal and team debriefing about progress and experience.

**Inclusion Criteria:** Include statements that refer to the implementation team’s (lack of) assessment of the progress toward and impact of implementation, as well as the interpretation of outcomes related to implementation. Reflecting and Evaluating is part of the implementation process; it likely ends when implementation activities end. It does not require goals be explicitly articulated; it can focus on descriptions of the current state with real-time judgment, though there may be an implied goal (e.g., we need to implement the innovation) when the implementation team discusses feedback in terms of adjustments needed to complete implementation. *(Any time team makes changes based on D2D results. Or when team gets feedback but comes up with reason to act or not.)*

**Exclusion Criteria:** Exclude statements related to the (lack of) alignment of implementation and innovation goals with larger organizational goals, as well as feedback to staff regarding those goals. Evidence of the integration of evaluation components used as part of “Reflecting and Evaluating” into on-going or sustained organizational structures and processes may be (double) coded to Goals and Feedback.

Exclude statements that capture reflecting and evaluating that participants may do during the interview, for example, related to the success of the implementation.

**NON-CFIR Codes**
| **A. Parallel Initiatives** | **Definition:** Simultaneously occurring intervention which share design characteristics with the intervention being implemented (and evaluated). For example, in the present study, a parallel initiative would be the participation in another Audit and Feedback program.

Inclusion Criteria: Include statements that refer to the participation in another Audit and Feedback Program and the description of such a program.

Exclusion Criteria: Exclude statements that compare one audit and feedback program to another, such as the advantages or disadvantages. Those should be coded under “Relative Advantage”. |