Computer Supported Cooperative Work (CSCW) (2020) 29:419–449
DOI 10.1007/s10606-019-09369-6

A Socio-Temporal Perspective on Pilot Implementation: Bootstrapping Preventive Care

Troels Mønsted[1]*, Morten Hertzum[2] & Jens Søndergaard[3]
*[1]*Department of People and Technology, Roskilde University, Universitetsvej 1 P.O. Box 260, 4000 Roskilde , Denmark (); [2]Department of Communication, University of Copenhagen, Karen Blixens Plads 8, 2300 Copenhagen , Denmark; [3]Research Unit of General Practice, Department of Public Health, University of Southern Denmark, J.B. Winsløws Vej 9A, 5000 Odense , Denmark*

Abstract. Systems for preventive care seek to provide healthcare services to citizens at risk of developing disease. In doing so they wrestle with identifying the citizens at risk of developing lifestyle-related disease and with reshaping the existing healthcare infrastructure into effective health offers for these citizens. In this study we analyze how a system for preventive care was enacted through a pilot implementation. The temporariness of the pilot implementation was, we argue, central to its contribution toward bootstrapping the system. By being temporary the pilot implementation became a means of acquiring clinical evidence for the cost-effectiveness of the system prior to committing to its long-term and large-scale use. The temporariness also legitimized temporary solutions to issues that otherwise made it difficult to bring actors, practices, and technologies into alignment. Once aligned, even if merely temporarily, the resurfacing of these issues after the pilot implementation will be shaped by the experiences from the pilot implementation. In this way pilot implementations have a generative role in infrastructure evolution; they are not merely tests but help bootstrap systems by making alignment manifest and benefits salient. We discuss this generative side of learning in pilot implementations and the extent to which they can enact a system by bootstrapping it.

Key Words: Alignment, Bootstrapping, Healthcare information systems, Pilot implementation, Preventive care, Temporary organizations

1. Introduction

Pilot implementations are a means to learn prior to the large-scale implementation of information systems (Hertzum et al., 2012). A pilot implementation is a temporary occasion for making it salient to a restricted group of users what working with a new system will be like. This real-use experience provides input to the finalization of the system, but the execution of the pilot implementation also marks agreement, at the management level, that the system has progressed sufficiently to warrant trying it out in real use. In this way a pilot implementation may start off a system by manifesting managerial alignment and by making the system salient to its users. The importance of such early experiences of alignment and real use lies in the otherwise slow pace at which infrastructures evolve (Bowker and star, 1999; Hanseth and Lyytinen, 2010). In this study we investigate the pilot implementation of a system that aims to evolve the Danish healthcare infrastructure by instituting an improved model for preventive care.

The slow pace at which infrastructures evolve tends to shift attention to extended timescales – years and decades. Thereby, design is turned into continuing design (Karasti et al., 2010), infrastructures into infrastructuring (Bowker and Star, 1999), and planned change into evolutionary dynamics (Hanseth and Lyytinen, 2010). Over extended periods of time the accumulated changes may transform the infrastructure but its reach – in terms of functional scope, technical interdependencies, and user base – makes the process open-ended and beyond the control of any individual actor (Ribes and Finholt, 2009; Star and Ruhleder, 1996). The extended timescale shows that the alignment demonstrated in a pilot implementation cannot be assumed permanent, but it may be sufficiently robust to warrant the large-scale implementation of the system. Furthermore, the deliberate temporariness of pilot implementations increases the possibilities for experimentation and innovation (Bakker et al., 2011; Scarbrough et al., 2004). Such possibilities are important to the incremental development of infrastructures because their interconnectedness tends to diminish the possibilities for experimentation. While diminished possibilities for experimentation stifle innovation, the details of how pilot implementations may enable experimentation and infrastructure evolution are not well-understood.

Because pilot implementations make a new system salient, although only to a restricted group of users, they compel these users to incorporate the system in their practices. This use ahead of large-scale use begins to tackle a key problem in infrastructure evolution. For the early users of a new system the inertia associated with the slow evolution of infrastructures often means that they are deprived of immediate benefits and compelling reasons to adopt. This tension between immediate benefits and long-term infrastructure evolution is known as the bootstrap problem: If it does not become beneficial to adopt a change until many others have already adopted it then the change is difficult to start off (Aanestad and Jensen, 2011; Hanseth and Lyytinen, 2010). With this study, we show that a pilot implementation can provide fertile ground not only for learning about the practical usefulness of a system, but also for bootstrapping it by providing a temporary framing that allows new alignments to emerge and settle.

Empirically, this study analyzes the pilot implementation of a system for the early detection and prevention of lifestyle-related disease. Preventive care consists of providing healthcare services to citizens at risk of developing disease (Glasgow et al., 2001). In support of this goal the system consisted of a stratification model for citizen screening and a cross-sectoral intervention involving general practitioners (GPs), municipal health professionals, and the citizens at risk. The pilot implementation was conducted to assess the feasibility of the system and to inform the further work on its design and implementation. To achieve its aim, the pilot implementation involved enrolling the various groups with a stake in the system and bringing their agendas into alignment. We analyze the role of the pilot implementation in bootstrapping the preventive care system. That is, we analyze (a) how a pilot implementation was recognized as an apt means for enacting the preventive care system, (b) how the pilot implementation was set in motion by bringing its actors into alignment, at least temporarily, and (c) how the learning from the pilot implementation was recoupled with the embedding healthcare infrastructure.

In the next section we account for related work on infrastructure evolution and pilot implementation. Then we describe the method of our empirical work, which is based on interviews, and present the findings of our analysis. Finally, we contribute a discussion of the ways in which learning happens in pilot implementations and the extent to which they can enact a system by bootstrapping it.

2. Related Work

The healthcare infrastructure is influenced by myriad socio-temporal issues (e.g., Zerubavel, 1979). Like other infrastructures it evolves through incremental innovation (Bowker and star, 1999; Hanseth and Lyytinen, 2010). In the following we review previous work on the evolution of infrastructures and on pilot implementation, which is an instance of temporary organization.

2.1. Infrastructure Evolution

The healthcare infrastructure has existed for centuries (Schatz and Berlin, 2011). During this period, it has been transformed beyond recognition. The scale of the transformation makes it important to remember that even large-scale infrastructures are 'enacted through the situated practices of specific community members in specific locations and time zones' (Orlikowski and Yates, 2002, p. 690). The enacted character of infrastructures denotes that they are worked out in the making, rather than developed and disseminated according to a preconceived plan. Most of the community members take the infrastructure for granted but they still contribute to enacting it through their daily practices. For example, any individual GP contributes to enacting the healthcare infrastructure by diagnosing patients, prescribing medication, having set consultation hours, being more or less empathic, and so forth. In contrast, the pilot implementation investigated in this study is an instance of a practice that specifically aims to change the infrastructure.

It is the enacted character of the studied preventive care system that motivates our socio-temporal perspective on its pilot implementation. To describe how infrastructures are enacted Jackson et al. (2011) emphasize their temporal structuring and distinguish among four kinds of rhythms: biographical, phenomenal, organizational, and infrastructural. The simultaneous presence of multiple, distinct rhythms creates tension and makes temporal alignment crucial to collaborative work (Chen et al., 2016; Jackson et al., 2011). For example, many citizens struggle to align their day-to-day exercise and diet habits (a biographical rhythm) with the long-term risk of developing a chronic lifestyle-related disease (a phenomenal rhythm). Relatedly, the studied pilot implementation aims to align its model for stratifying citizens into risk groups (an organizational rhythm) with the national agreement about how GPs are reimbursed for their work (an infrastructural rhythm). Ensuring that design and development activities contribute value to the long-term cultivation of an infrastructure involves a diverse set of temporal alignments. Karasti et al. (2010) emphasize one of them, namely the need to align infrastructure time with project time. Infrastructure time is the temporal orientation held by the users who are accountable for the long-term effects and usefulness of an infrastructure. Their primary concern is the gradual evolution of the infrastructure across a series of projects. This temporal orientation values developments that are flexible, useful in practice, and made with a view to later-and-larger changes. In contrast, project time is the temporal orientation held by the designers involved in shaping and building infrastructure components. These designers are primarily concerned with delivering technically well-designed components on time. This temporal orientation values fixed component specifications and tends toward a here-and-now focus.

The interactions between infrastructure time and project time revolve around an installed base of preexisting practices, standards, IT capabilities, and other sociotechnical elements. While an infrastructure develops through the evolution of the installed base (Aanestad et al., 2017), the relation between the infrastructure and the installed base is one of both tension and opportunity. In the words of Star and Ruhleder (1996, p. 113) an infrastructure 'wrestles with the "inertia of the installed base" and inherits strengths and limitations from that base'. While the installed base contributes welcome longevity, it is an important limitation that the infrastructure also becomes slow-changing, possibly to the point of lacking the flexibility necessary to grow and expand in response to new demands.

The lack of flexibility includes the risk that extensions of the infrastructure may not gain momentum because early users perceive no immediate benefit of adopting them. Hanseth and Lyytinen (2010) refer to this risk as 'the bootstrap problem'. It is especially large for extensions that gain their value from having many users and, thus, provide little incentive to early adopters (e.g., the first few phone owners could merely call each other). To counter the bootstrap problem an infrastructure extension must (a) be directly useful, even without a large user base, (b) exploit the existing installed base so that there is little or no need for additional support, and (c) make the user base grow at a rate sufficient to sustain the cost of development and learning (Hanseth and Lyytinen, 2010). Relatedly, other studies propose that enacting an infrastructure requires aligning stakeholder goals to motivate contribution and enroll stakeholders in infrastructure use (e.g., Randell et al., 2015; Ribes and Finholt, 2009).

While the bootstrap problem concerns the adoption of an infrastructure extension and thus its transition into use, research on the evolution of infrastructures tends to avoid a distinction between design and use phases. For example, Karasti et al. (2010) suggest continuing design as a development approach that recognizes infrastructure time. Continuing design emphasizes continuity to ensure that a working system is always in place. And it blurs the boundaries between design, implementation, use, and maintenance by emphasizing incremental change and deemphasizing project-time thinking. Similarly, Pipek and Wulf (2009) propose that an integrated perspective on design and use will improve organizations' efforts to evolve their infrastructures. The principal reason for avoiding a distinction between design and use phases is that infrastructures evolve without a well-defined end point (Aanestad and Jensen, 2011; Hanseth and Lyytinen, 2010). Over time, new uses emerge in response to new user needs, new technological capabilities, new organizational responsibilities, new international standards, and so forth. These emergent uses call for evolving the infrastructure and may, in turn, cause additional uses to emerge. In spite of this intractability Steinhardt and Jackson (2014) find that plans play an important role in infrastructures by facilitating the alignment of rhythms. Plans play their role partly when they are made and partly when they are acted out. When made, plans contribute to anticipation work, that is, to setting a forward-looking frame of reference, which among other things devises pathways from the present to an imagined future state of affairs (Steinhardt and Jackson, 2015). When acted out, plans are resources for situated action (Suchman, 2007). Though a plan does not determine the course of action, it may contribute to maintaining collaborative agreement about alignment – or more broadly about visions – in the face of dynamic circumstances (Steinhardt and Jackson, 2014, 2015).

2.2. Pilot Implementations

While the present study investigates the role of pilot implementations in bootstrapping an extension of an infrastructure, pilot implementations are most commonly depicted as a method for improving the quality of an information system and reducing implementation risk (Hertzum et al., 2012). In contrast to prototype tests in the laboratory pilot implementations are conducted in the field and, thus, sensitive to the embedding infrastructure. To be tested in the field a system must be properly engineered, but during a pilot implementation the system has not yet been finalized. Rather, the experiences from the pilot implementation are fed back into the finalization of the system. The role of pilot implementations in bootstrapping stems from the alignment that must be reached in preparing them and from their introduction of use ahead of large-scale use.

A pilot implementation is not just the period during which an information system is in pilot use. Hertzum et al. (2012) propose that pilot implementations consist of five activities: planning and design, technical configuration, organizational adaptation, pilot use, and learning. The three first activities are preparations. During the preparations the focus and scope of the pilot implementation are defined, the system is configured for the pilot site, operational data are migrated to the system, interfaces to other systems are established, work procedures at the pilot site are aligned with the system, users receive training, safeguards against breakdowns are set up, and so forth. The preparations may consume more time than the period of pilot use, during which the staff at the pilot site uses the system for real work. Finally, learning about the system, its implementation, and use occurs during the preparations as well as during the period of pilot use.

Pilot implementations are, by definition, temporary (Hertzum et al., 2012). The temporariness is essential because it creates a decision point: what happens after the pilot implementation? That is, the temporariness helps maintain that pilot implementations, unlike conventional implementations, are performed to inform decisions about how and, ultimately, whether to finalize and implement the system. Previous research shows that the learning objective of pilot implementations is often difficult to fulfill, for example because the different stakeholders in a pilot implementation may perceive its outcome differently (Winthereik, 2010). However, other literature suggests that the temporariness of pilot implementations creates room for experimentation and innovation. To investigate this, we draw on the literature on temporary organizations. In this literature 'temporary' refers to predetermined duration, not necessarily short duration (Bakker et al., 2016).

Lundin and Söderholm (1995) divide the lifecycle of a temporary organization into four phases and discuss the boundary work associated with each phase. During the first phase an entrepreneur *makes the case for creating a temporary boundary* around a specified activity and, thereby, setting it apart from the embedding organization. Making a case involves, among other things, establishing synchronicity among multiple agendas (Granqvist and Gustafsson, 2016). For a pilot implementation this phase ends with the decision to conduct the pilot implementation. During the second phase the temporary organization is planned and set up. This phase involves boundary work to *decouple the temporary organization by bracketing it from the embedding organization*. It corresponds to the three preparation activities in the model of pilot implementations by Hertzum et al. (2012). The third phase, termed planned isolation by Lundin and Söderholm (1995), corresponds to the period of pilot use. In this phase work goes into *guarding the boundary* to protect the temporary organization from disturbances as well as into *maintaining contacts across the boundary* to exchange information. The fourth phase is the termination of the temporary organization. It involves boundary work to *recouple the temporary organization with the embedding organization*, including the transfer of knowledge and experience gained in the temporary organization. Like decoupling, the recoupling phase can be seen as a process of aligning organizational rhythms with infrastructural rhythms (Jackson et al., 2011).

One of the main features of temporary organizations is that their isolation from the embedding organization creates room for innovation and learning by shielding activities from the sanctions to which they would otherwise be exposed (Bakker et al., 2011; Scarbrough et al., 2004; Zietsma and Lawrence, 2010). That is, a pilot implementation has more degrees of freedom to experiment than can be expected in other activities, which to a larger extent are constrained by the interconnectedness and inertia of infrastructures. At the same time the temporariness seems to inhibit the sedimentation of the resulting knowledge because when the temporary organization terminates and its members move on, the knowledge is likely to disperse. This creates what has been termed a learning paradox: While the temporary organization is an effective vehicle for knowledge creation, it can substantially impede knowledge sharing (Bakker et al., 2011). With respect to pilot implementation the learning paradox may play out quite differently for the learning that occurs during the preparations and the learning that occurs during the period of pilot use. The former occurs before the isolation phase and, thus, at a point in time where the pilot implementation and the embedding organization are, at least partially, learning together. The latter occurs during the isolation phase and is, thus, subject to the full impediment of knowledge sharing. This is reminiscent of Brown and Duguid's (2001) argument that knowledge is transferred through shared practices and, thereby, that it is at divisions in practice that knowledge stops flowing.

The temporary organization is not completely shielded from the embedding organization. Its isolation is negotiated and permeable, not absolute. The permeability means that the work performed by the actors in a temporary organization is influenced by the shadow of the past and the shadow of the future (Grabher, 2002; Ligthart et al., 2016). While the shadow of the past is the actors' experiences with each other from interactions prior to the temporary organization, the shadow of the future is their expectations about further interactions after the conclusion of the temporary organization. These concepts reach across decoupling and recoupling and, thereby, acknowledge the inescapable interconnectedness of infrastructures. It is exactly because decoupling and recoupling are partial and negotiated that work is required to maintain the boundary between the pilot implementation and the embedding infrastructure (Stjerne and Svejenova, 2016). However, the literature on temporary organizations is, in general, more open to the possibility of temporarily bracketing at least some of the interconnections than that on infrastructures.

The work involved in, partially, decoupling and subsequently recoupling a pilot implementation and the embedding infrastructure is essential to the possibilities and limitations of pilot implementation. For example, Hertzum et al. (2019) find that the learning from pilot implementations is situated and messy because it is difficult to tell the temporary consequences that are particular to the pilot implementation from the lasting consequences of the system once it is fully implemented. The reason for this difficulty is ambiguity about decoupling and recoupling; the consequence is uncertainty and, possibly, confusion about what can be learned from the pilot implementation. Even the decision about when to decouple and recouple is likely to cause debate. With respect to decoupling, Buxton (1987) has noted that 'it is always too early (for rigorous evaluation) until, unfortunately, it's suddenly too late.' With respect to recoupling, pilot implementations should be sufficiently long to reach a level of stable use and allow for realistic assessment, but at the same time pilot implementations should be short because 'full-scale implementation awaits completion of [the] pilot' (Pal et al., 2008, p. 261).

3. Method

The empirical context for this study was the project TOF ('Tidlig Opsporing og Forebyggelse', Danish for Early Detection and Prevention). TOF developed and pilot implemented a new preventive care system. We analyze the role of the pilot implementation in bootstrapping this system.

3.1. Setting: The TOF Project

TOF was initiated in 2009 on the initiative of the Region of Southern Denmark (RSD) with the ambition of developing an improved model for preventive care. Most previous projects had relied on a preventive care model where all citizens (within a certain age range) were invited for a health interview with their GP. TOF was to produce the same health benefits for the citizens with less effort from the health professionals. Fulfilling this ambition required an accurate means of detecting the citizens at risk of developing lifestyle-related disease and a more efficient utilization of existing preventive health offers. The development of an information system was considered pivotal to meeting these two requirements. First, this system should automatically stratify citizens into risk groups on the basis of data self-reported by the citizens and data retrieved from the medical records of the citizens' GP. Second, citizens at risk should be channeled to either general practice or the municipal health service for the health offer best suited to their risk group. During the first years of TOF, the components of the system were conceptualized and developed by the Research Unit of General Practice (RUGP) at the University of Southern Denmark in collaboration with two end-user working groups: the GP working group and the municipal working group.

In 2012, a clinical feasibility study of the system was conducted with four GP clinics. The positive outcome of the feasibility study secured funding for the pilot implementation that is the focus of this study. However, the preparations for the pilot implementation continued from 2013 to 2016, primarily due to two contingent events. We will elaborate on these events in the analysis but resolving them entailed extensive negotiations with, among others, RSD, the Organization of General Practitioners in Denmark (OGP), and the ten municipalities taking part in TOF. The period of pilot use finally took place in the fall of 2016. It lasted three months and saw the participation of two municipalities, 47 GPs from 18 clinics, and 2661 citizens (out of 9400 invited).

In the fall of 2017 (i.e., the time of this writing) the evaluation of the pilot implementation concluded that the TOF model for preventive care was promising but not yet ready for large-scale implementation. In particular, the stratification model for automatically dividing the citizens into risk groups had to be improved. It was decided that the next step of TOF would be a second pilot implementation.

3.2. Procedure

We collected data about the pilot implementation by means of interviews conducted shortly after the period of pilot use (January–February, 2017), supplemented with analysis of project documents. The choice of interviews and document analysis was a result of the retrospective nature of the study. Our study protocol and interview guide were presented to the TOF research steering group. Informed consent was obtained from each interviewee, including permission to audio-record the interviews.

We conducted a total of 60 interviews for this study. The interviewees were purposefully sampled to combine near complete coverage of the people involved in establishing and conducting the pilot implementation with representatives of the citizens and health professionals involved in the period of pilot use. In total, the sample of interviewees spanned three levels of involvement in TOF: the project level, the practice level, and the regional/national level. The project level covered the people involved in developing the information system, negotiating the associated health offers, and organizing the pilot implementation. At this level we interviewed the head of the TOF project (RUGP), the IT project manager (RUGP), the research coordinator (RUGP), a graphic designer (RUGP), a researcher affiliated to the project (RUGP), two project managers (RSD), and two internal consultants on general practice (RSD). The practice level covered the people directly involved in the use of the system. We interviewed seven GPs (from seven different clinics in two municipalities), one general-practice nurse from one of the clinics, four municipal health professionals (two from each municipality), and fifteen citizens, of which thirteen were interviewed twice (during and after their participation in the pilot implementation). Finally, the regional/national level covered organizations with various labor-market, political, and healthcare interests in TOF. At this level we interviewed four representatives from OGP (including the current and former head of OGP division south), three from RSD (the former director of health, the head of the department of cross-sectoral collaboration, and the former head of the department of research and quality), two from patient associations, and two from national health offers directed at lifestyle-related health issues (drinking and smoking).

The interview guide revolved around (a) the interviewees' involvement in TOF, (b) their knowledge of the process leading up to the pilot implementation, (c) their experience of the TOF system during the pilot implementation, and (d) their perception of what was learned from the pilot implementation. All four topics were addressed in all interviews but to different extents and in somewhat different ways depending on the interviewee's role in relation to TOF. For example, the process leading up to the pilot implementation was covered in more detail during interviews at the project and regional/national levels. At the practice level, this topic mostly concerned how, and why, the interviewee got involved in the pilot implementation. For practical reasons 36 of the interviews were conducted over the phone; the other 26 interviews were conducted face to face at a location convenient to the interviewee. The majority of interviews lasted 45–60 min, but some of the follow-up interviews with citizens lasted 15–30 min.

3.3. Data Analysis

The audio recordings of the interviews were partially transcribed. That is, we listened through the audio recordings and transcribed the parts about how the TOF system was enacted through the pilot implementation. Listening through the recordings served to become sensitized to the content of the interviews; the transcripts provided a record of this listening process by including verbatim interview passages as well as descriptive annotations. The data analysis proceeded by iteratively reading and coding the transcripts in a ground-up manner consisting of open coding followed by axial coding (see Strauss and Corbin, 1998). In the beginning the codes were phrases used by the interviewees or, in the absence thereof, simple descriptive phrases. This open coding produced a number of meaning-bearing phrases, or concepts, and was followed by a search for linkages between the concepts. Through this search we gradually identified axes around which many of the concepts could be organized. The axial coding grouped the concepts into themes of interlinked interview content. Identifying the axes also meant becoming aware of new concepts in the data. These new concepts were added to the open coding and incorporated in the continued axial coding. Relatedly, some concepts that appeared important early in the analysis faded into the background and were eventually dropped.

The themes identified through the axial coding were distributed over the period from 2009 to 2017. In the final stage of the coding we arranged the themes into three temporal phases that concerned (a) how the pilot implementation became recognized as an apt means for enacting the TOF system, (b) how the pilot implementation was set in motion by bringing its actors into alignment, and (c) how the learning from the pilot implementation was recoupled with the embedding healthcare infrastructure. Figure 1 shows the structure of concepts, themes, and phases that emerged from the data analysis, thereby summarizing how the data were coded. In addition, the figure aims to provide an up-front overview of our main findings.

4. Findings

The initial intention of RUGP was to develop the TOF system as a small-scale project to enable incremental development and documentation of its components and health outcomes. However, it soon grew into a large initiative. There were several reasons for this growth. First, the project early demonstrated its potential for improving preventive care and thereby caught the attention of RSD and other stakeholders because disease prevention was high on the political agenda. Second, to produce scientific knowledge about the health outcomes of TOF it was necessary to conduct a clinical trial on a large population of citizens. As a consequence, RSD decided to allocate a special grant to TOF, which in response grew in magnitude and infrastructural complexity.

In the following, we present our analysis of how the pilot implementation contributed to the bootstrapping of the TOF system. We start with the three themes in the phase that led to the decision of conducting the pilot implementation. Subsequently, we get to the two other phases of Fig. 1.

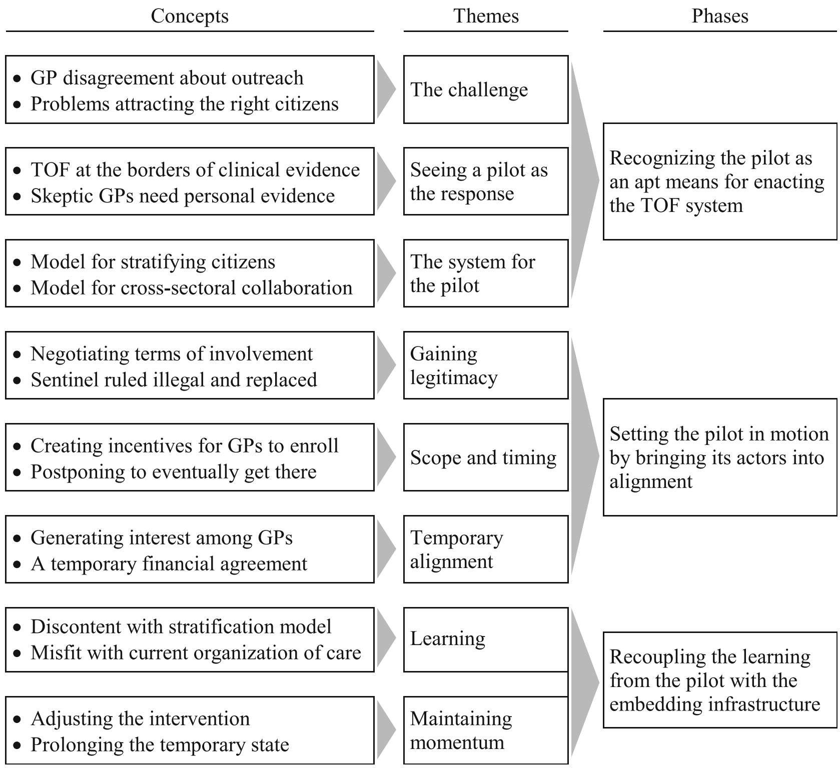

Fig. 1. The structure of concepts, themes, and phases that emerged from the data analysis

4.1. The Challenge

The TOF project was introduced into an installed base of preventive health offers. Two actors were particularly prominent in providing these offers or channeling citizens to them: the GPs and the municipal health centers. Preventive care was, however, a debated and complex issue. Among the GPs there was widespread disagreement about their role in relation to citizens with a behavior that incurred risk of lifestyle-related disease. While the obligation and importance of general practice to engage in preventive care was not questioned, the GPs disagreed about their role in reaching out to citizens. Some GPs were strongly in favor of outreach activities:

Thirty years ago, when I started as a GP, the general idea was that citizens were responsible for their own health, and we should only treat them when they fell ill. Today, we get paid for reaching out; it has simply become part of the reimbursement system. (GP 1).

Other GPs were concerned about the potential adverse effects of preventive care. They referred to studies of pathologization. These studies indicated that when citizens with lifestyle issues entered the healthcare system and became subject to medical scrutiny then the outcome was sometimes increased depression and suicide rates. These GPs argued that the risk of such adverse effects of TOF was unknown:

By checking the citizens' lung function, taking an ECG [electrocardiogram], and inviting them for a health interview, you do something to them. If you do this, you will need to have a strong expectation that it will be to their benefit, but who knows how many develop a depression or are driven toward suicide. We don't know. (Head of OGP division south).

The municipal health centers had been established (following a reform of local government in 2007) with the aim of providing a centralized entry point to preventive health offers, such as alcohol treatment, dietary consultancy, and smoke-cessation courses. While these health offers were generally considered to be of good quality, they struggled to attract the target citizens in sufficient numbers. Thus, the municipalities involved in TOF saw it as an important, but unresolved, issue to find ways of reaching out to a larger part of the target population:

In the Region of Southern Denmark, the municipalities have put a real effort into providing relevant offers for the citizens, but it is a challenge to make enough citizens use these offers. In reality, we end up doing something good for very small groups of citizens. (Internal consultant on general practice 1, RSD).

The challenge of attracting the citizens at risk of developing lifestyle-related disease was also well-known to the GPs. For example, one of the interviewed GPs stated that part of her motivation for taking part in TOF was to 'keep the healthy marathon runner out of the clinic and get to see those who are otherwise just sitting down at the local pub'. While the former group was more attentive to health offers, the GP felt that she had more to offer to the latter group.

The debated and unresolved character of preventive care meant that the relations between TOF and the installed base of preventive health offers were not clear from the outset. Rather, TOF had to work with the GPs, their organization OGP, the municipalities, and the healthcare region RSD to bring these key stakeholders into alignment. To succeed in this process TOF must be able to accommodate their requirements as they gradually took form and to keep the stakeholders on board by providing initial results that could fuel the long-term process.

4.2. Seeing a Pilot Implementation as the Response

The development of the TOF system took place from 2010 onward with the model for stratifying citizens into risk groups as a main activity. To ensure the clinical validity of the stratification model, RUGP made an initial version of the model on the basis of scientifically validated scores for identifying selected lifestyle-related diseases. In parallel, the municipal and GP working groups were formed to fit the stratification model to current practice and define the details of the intervention. Finally, a feasibility study was conducted to test the clinical precision of the stratification model. Through the development and testing of the stratification model it became increasingly clear that the project operated on the borders of existing clinical evidence. This realization gradually shifted the focus of TOF from an implementation and operation project to a research and innovation initiative:

Somewhere along the way, the project changed. To begin with it was thought of as an operation project run by RSD with the assistance of RUGP, which would mainly be responsible for the research evaluation including the coordination of the research units involved. For several reasons this framing changed [and TOF increasingly became a research project]. (Research coordinator, RUGP).

A major driver in this shift was the project stakeholders' inclination to evidence-based thinking. In the absence of existing clinical knowledge, it fell on the TOF project to provide evidence for the validity and precision of the stratification model. This need for evidence became an important enabling argument for the pilot implementation but it also placed high demands on the pilot implementation and thus its size:

RSD thought that it [the pilot implementation] was too small. Also, we saw that if we were to test whether the concept would work in reality then one municipality would not be enough. (Researcher affiliated to TOF, RUGP).

Another enabling argument for the pilot implementation was persistent skepticism, in particular among the GPs, about the effectiveness of the kind of health checks proposed. This skepticism made it insufficient to specify the content of the health checks in detail; the GPs needed practical experience with conducting the health checks to discover for themselves whether their skepticism was warranted. This argument for the pilot implementation was less about clinical evidence and more about personal evidence in terms of an opportunity to experiment with how to conduct health checks and experience first-hand whether they had an impact on the citizens:

There are many previous research studies on health checks, and the overall conclusion must be that they didn't do much good. But if the citizen has a high-risk profile then they should work, provided that we do our job well and actually move the citizen. (GP 2).

In view of these concerns, OGP made their support of the project subject to a thorough evaluation of the effects of TOF before a larger implementation. That is, at this stage a pilot implementation had become a demand from a key stakeholder in the project. In the same period, the information system necessary to support the project grew in scope in order to acquire the input for the stratification model, communicate the outcome of the stratification to citizens and health professionals, and so forth. With a growing information system came an increased need for evaluating whether it performed its part in the intervention:

The IT support has grown larger. For instance, […]: How are the citizens informed? Do they understand the informed-consent form they sign [electronically]? How do they understand the feedback they receive through the system? These were the most important points that a pilot implementation could help us understand better. (Research coordinator, RUGP).

The final closure to these considerations came in 2013, when RSD decided to fund the development and implementation of TOF through a €5,000,000 grant, which at the time was RSD's largest investment in health research. The investment was specifically earmarked to a pilot implementation that should enable the project to generate the knowledge necessary to decide whether the TOF system should be permanently implemented in the region.

4.3. The System for the Pilot Implementation

Following the investment from RSD the TOF system was made ready for the pilot implementation. The system consisted of the stratification model, an organizational model, and an information system.

First, the *stratification model* categorized the citizens into high risk, moderate risk, low risk, and those with a preexisting diagnosis. This stratification was made using data from two sources. The first data source was a 15-item questionnaire filled out by the citizens. Questionnaire items were based on the Swedish National Guidelines for Disease Prevention and covered the citizens' alcohol consumption, diet (intake of sweets, fish, fruit, vegetables), physical exercise, smoking habits, observable symptoms (coughing, shortness of breath), and perceived general health. The second source of data about the citizens was clinical data extracted from their GP's electronic medical records. These data included prescription codes, National Health Service disbursement codes, and International Classification of Primary Care codes. On the basis of these data the citizens at high risk of developing lifestyle-related disease were identified using three validated risk scores (the COPD-PS screener, the Danish Diabetes Risk Model, and the Heart-score). The citizens stratified to be at moderate risk were those with an unhealthy lifestyle, defined as an alcohol intake exceeding 14 (females) or 21 (males) units per week, a body mass index exceeding 30, a score of less than 4 on the Swedish National Guidelines for Disease Prevention, or less than 150 min of physical activity per week. Citizens who did not meet any of the abovementioned criteria were stratified as low risk, and those who already had a diagnosis were stratified as such.

Second, a novel *organizational model* was devised to make full use of the existing preventive health offers in general practice and the municipalities. The ambition was to transform these offers into a coherent preventive intervention by establishing cross-sectoral collaboration among the involved care providers. In the organizational model the citizens at high risk were encouraged to make an appointment with their GP. For those who did, the GP was responsible for offering a thorough health examination, including a broad spectrum of blood tests, and a health interview for discussing the citizen's risk profile and care needs. Based on the outcome of the interview the GP was, furthermore, responsible for offering the citizen follow-ups (e.g., a monitored weight-loss program) or for referring the citizen to relevant external offers (e.g., in the municipality). In contrast, the municipalities were assigned responsibility for the citizens at moderate risk. The municipal health centers were to contact all these citizens by phone to conduct an initial health interview about the citizen's risk profile and motivation for behavior change. Based on the outcome of this interview, the municipal health center could either offer an extended meeting in person or refer the citizen to municipal health offers such as dietary consultancy or a smoke-cessation course. Finally, the municipal health center might advise the citizen to attend care at the GP. While the GPs and municipal health centers were mainly responsible for citizens in different risk groups, they were also responsible for referring citizens to further health offers in general practice or the municipality. Such cross-sectoral referrals required that the GPs were sufficiently aware of municipal health offers, and vice versa.

Third, an *information system* was needed to apply the stratification model and organizational model efficiently in practice. One facility of this information system was a web-based health folder for communication with the citizens, GPs, and municipal health centers. The citizens filled out the 15-item questionnaire in the health folder and received the result of the stratification. Specifically, citizens stratified to be at high risk were encouraged to schedule an appointment with their GP and citizens at moderate risk were informed that they would be contacted by a municipal health professional. Citizens at low risk and those with a preexisting diagnosis received the result of the stratification and were informed that this concluded their participation in TOF. For the GPs and municipal health centers the health folder presented details on the citizens' health condition and risk behavior. Another important facility in the information system was the data-capture tool that extracted data from the GPs' electronic medical records for use in the stratification model.

4.4. Gaining Legitimacy

In 2014 the information system was ready for pilot implementation, and TOF proceeded to the second phase in Fig. 1: setting the pilot implementation in motion. This phase involved establishing the temporary organization that would allow for testing the TOF system in practice. To establish the temporary organization TOF needed to enroll the necessary number of municipalities and GPs. A key element in negotiating the terms of their involvement was to maneuver TOF into a legitimate position.

For the municipalities, participation was based on voluntarism, which was deemed viable due to their great interest in strengthening preventive health offers. Three of the ten municipalities in the project initially volunteered to participate in the pilot implementation but one municipality later withdrew out of concerns that its shortage of GPs did not leave resources for its participation. However, the two remaining municipalities provided a sufficiently large population for the pilot implementation to yield the required insights about the stratification model and the rest of the TOF system. For the GPs, RSD initially pushed for making participation mandatory for all GPs in the participating municipalities to ensure that the intervention would be available to as many citizens as possible:

From RSD's perspective it was important that the agreements that we reached covered all GPs. For the municipalities it was also important to ensure participation from as many GPs as possible, so you don't end up in a situation where the offer is only available for 10% of the citizens. (Project manager 1, RSD).

During the negotiations it became clear that this approach would likely meet resistance from many GPs. There was a shortage of GPs in the region due to a wave of retirements and difficulties attracting new GPs, especially in rural districts. For that reason, many GP clinics ran at maximum capacity and felt an obligation to devote their time to patients with acute health conditions rather than to participation in innovation projects:

The GPs will hold the time that they spend up against the number of patients we get in, to assess whether or not it makes sense. We are very focused on time these years because we are so insanely busy. (Head of OGP).

Without a mandate from its members to support mandatory participation OGP insisted that GP participation was made voluntary. Due to OGP's powerful position as the organization that handled the interests of the GPs at the political level, RSD saw no alternative but to accept voluntary participation. Otherwise, the pilot implementation would, probably, have ground to a halt. By redefining the terms of participation TOF gained sufficient legitimacy to proceed.

In the fall of 2014, the legitimacy of TOF was however endangered by a contingent event. To ensure the robustness and scalability of the information system it had been decided to base the extraction of data from the GPs' medical records on two systems in the installed base: the data capture tool *Sentinel* and the database *DAMD*. Sentinel and DAMD were already in mandatory use in all GP clinics to produce quality reports for the individual clinics. In addition to being readily available these two tools were developed by the software partner in TOF and they could easily be extended to extract data for use in the stratification model. There were, however, growing concerns among GPs about potential misuses of the data extracted by Sentinel. Most notably, it was seen as a potential control mechanism because it in principle, but in sharp contrast to its actual use, could be used to benchmark GP clinics and thereby give RSD an advantage in financial negotiations:

So the general opinion about Sentinel changes from seeing it as something we have developed to support the development out in the clinics, to seeing Sentinel as the region's control tool. It was weird to witness such a change in public opinion among the GPs. (Researcher affiliated to TOF, RUGP).

The discontent with Sentinel was fully ignited in September 2014, when a major Danish news media reported that the data extracted with Sentinel and recorded in DAMD were allegedly not legal. The core of the allegation was that the data in DAMD covered more diseases than the database had been approved for by the authorities. This caused a lengthy political and legal struggle that eventually led to the conclusion that some of the data were not legal; in the spring of 2015, DAMD was deleted. While Sentinel was only indirectly involved in this conflict, it increased OGP's concerns about potential misuses of data. As a direct consequence OGP decided to suspend their support of Sentinel until an independent evaluation had determined whether it handled data in a secure manner. This decision effectively blocked the use of Sentinel in TOF and threatened the legitimacy of the project because it was now associated with data misuse. To regain legitimacy, the project first decided to wait for the independent evaluation of Sentinel, hoping that it would be quick and enable the project to proceed with its original strategy. This strategy however ran counter to the interests of RSD:

There were many critical steering-committee meetings in 2015 that, in principle, could have ended the project. There was internal disunity because some representatives from RSD wanted the project to continue quickly, while you could sense that the mandate of the OGP representatives was to draw the project in a different direction. That resulted in some steering-committee meetings with loud arguments. (Research coordinator, RUGP).

During the spring of 2015 it became increasingly apparent that TOF needed another strategy for data extraction because OGP's reassessment of Sentinel proved to be a lengthy affair. The solution came when TOF proposed a data-extraction approach that gave citizens and GPs authority over the data exchange, in contrast to the bulk extraction in Sentinel. Specifically, it was proposed that the citizens should give explicit consent to the data extraction, that the GPs should individually instruct their system provider about which data to extract, and that this process should be supported by a new information system developed from the ground. Bypassing the Sentinel and DAMD systems in the installed base delayed the project but was necessary to regain legitimacy. OGP accepted the new approach in late spring 2015 and TOF could proceed with the pilot implementation.

4.5. Scope and Timing

TOF successfully avoided the threats against the legitimacy of the project by adjusting some of its key components but these adjustments influenced the scope and timing of the pilot implementation. While mandatory GP participation would by definition secure a large user base, the shift to voluntary participation necessitated personal incentives for the GPs to participate. To yield robust learning about the health effects and organizational feasibility of TOF it was assessed that at least 50% of the GPs in the two municipalities needed to enroll. A prerequisite for attaining this level of enrollment was that RSD and OGP could reach a financial agreement that would ensure the GPs an attractive reimbursement for their participation. Because healthcare in Denmark is publicly funded but the GPs are private agents, almost all activities in general practice are fully reimbursed through disbursement codes. These codes are specified through a common agreement negotiated between the regions and the OGP at the national level. If an activity is not covered by the common agreement, it is necessary to negotiate an addendum, referred to as a *local agreement*. It is often the case that innovation projects involve such local agreements.

Under normal circumstances this would have been a routine negotiation involving an assessment of the time required for the GP to perform the health examination and health interview involved in TOF; these activities would then be reimbursed at a standard rate. At the time of the negotiation of the local agreement for TOF the negotiation climate between RSD and OGP was, however, in a state of crisis. This contingency was the result of lengthy but failed attempts to seal the common agreement (renegotiated every third year). Eventually a government intervention was launched to seal the common agreement. In this climate of conflict and mistrust among the parties, progress on the local agreement for TOF was effectively blocked. The situation was further complicated for TOF because data security was an important item on the political agenda after the DAMD case:

The project is about early detection and prevention, but because the project originally made use of automated data exchange, the DAMD case became a major issue. The project was made an example of debates it was not directly part of. It is my impression that the project was used as an example by OGP to demonstrate to the region, that they cannot just dictate what projects the GPs should engage in. (Project manager 1, RSD).

Because TOF was critically dependent on the local agreement, the project had either to accept that the political climate blocked any possibility of enrolling GPs or to postpone the pilot implementation in the hope that RSD and OGP would reach common ground at a later point. In the steering committee it was decided to wait it out and instead spend the time improving specific components of the TOF system, in particular the interface of the health folder. In the fall of 2014 the local agreement was finally sealed and the enrolment of GPs could be initiated. While these events delayed the pilot implementation by at least a year, the willingness of TOF to adjust its scope and timing proved essential to aligning the project with the embedding healthcare infrastructure.

4.6. Temporary Alignment

The GPs found the local agreement satisfactory because the payment was assessed to fully match the actual time they would spend conducting health examinations and health interviews:

I believe that we reached a very good local agreement and that the GPs are satisfied with it because they get the money they should for the job they do. And when the agreement is right there will be little resistance outside the political system. Then people just do their job and get paid for it. This is not rocket science, that is what we do every day. (GP 2).

However, the events surrounding the DAMD case showed that the obtained alignment was merely temporary. The replacement of Sentinel with another tool for data extraction necessitated a revision of the local agreement for TOF because the new tool was more time consuming for the GPs to use. At that time, the negotiation climate between RSD and OGP had fortunately improved, and the revision of the local agreement in late summer 2015 went smoothly. The project could therefore begin to communicate with the GP clinics about their enrollment. In particular, a project manager from RSD was given the responsibility of visiting clinics and contacting GPs by phone. This approach generated a lot of interest:

When I began to call the GPs I soon learned that they in reality weren't very concerned about the research agenda in the project. They wanted to know whether we [RSD] backed the project and what it practically meant for them, what to do and when. In the situation we were in, it was really nice to experience that we could take the lead, make a lot of phone calls, establish good relations with the personnel in the clinics, and make them comfortable about the project. It took a lot of time, but also meant that we succeeded in recruiting two thirds of the GPs rather than the 50 % that were our target. (Project manager 1, RSD).

The adjustments made to the TOF system during the period of crisis combined with their ability to create incentives and communicate these to the GPs were decisive for the establishment of the temporary organization required to carry out the pilot implementation. The 47 (69%) GPs who enrolled exceeded expectations and provided the basis for a solid assessment of the stratification and organizational models. However, the adjustments also created an organization that was explicitly temporary because it was enabled by a local agreement that was an exception from the common agreement. While the common agreement was part of the lasting organization of preventive care, the local agreement was external to it and valid for the pilot implementation only. To ensure the sustainability of the TOF system and enable it to be scaled up, it would later be necessary to get it funded by negotiating its inclusion in the common agreement.

4.7. Learning

The period of pilot use finally started in September 2016 and continued for the planned three months. Of the 2661 participating citizens 22% were stratified to have high risk of developing lifestyle-related disease, 23% to have moderate risk, and 29% to have low risk. The remaining 26% were already diagnosed with a lifestyle-related disease. In this third phase, the TOF project worked to recouple the learning from the pilot implementation with the embedding healthcare infrastructure.

On the positive side, the pilot implementation demonstrated the practical and technical feasibility of the TOF system. The GPs were generally satisfied with the reimbursement model, because they found that the compensation matched the time they spent on the intervention. The information system proved robust and enabled the extraction and analysis of the health data. The usability of the health folder was satisfactory, although the diverse population of citizens created some problems. In particular, citizens with low health competence experienced barriers in understanding the health-related terms and formulations. Finally, the citizens, who self-selected to participate, tended to appreciate the opportunity to have their health checked:

I was a bit surprised to be in the category that needed a check. But that may just confirm that it was a good idea to be checked. (Citizen 1).

The learning from the pilot implementation was, however, also that several obstacles had to be addressed before the TOF system was ready to be extended to more municipalities and included among their permanent health offers. One of these obstacles was that the thresholds in the stratification model were not well-aligned with the local organization of preventive care. The GPs and municipal health professionals perceived the citizens to be less at risk of developing lifestyle-related disease than the model predicted. Specifically, the majority of the GPs reported that most of the citizens they received either did not qualify as being at high risk or were citizens with a sound understanding of their health issues:

I think far too many citizens have been referred [to their GP]. The threshold is set too low. Some have been sent here just because they are slightly overweight or have slightly unhealthy eating habits. I believe they should just have been sent straight to the municipal health offers. It seems as if there has been no selection at all. All seem to have been uncritically stratified as high risk. (GP 3).

Hence, many GPs found that they did not have relevant offers to the citizens and that they should, instead, have been referred to the municipal health centers for dietary consultancy or the like. That is, the citizens stratified to be at high risk were, in many cases, perceived to be merely at moderate risk. This view was shared by many municipal health professionals, who would also have preferred to spend their efforts on citizens with more urgent needs. For the patients stratified to be at high risk it was confusing to experience a mismatch between their stratification and their GP's recommendations. These citizens tended to struggle to make sense of the experience, often concluding that it had been pointless:

There is nothing alarming in the tests but the questions I have answered beforehand trigger that my diet needs changing. However, I am not overweight and she [the GP] does not think I need to change anything. That is also how I feel. I am not feeling bad or in pain or that I am in poor health. I came because I am taking part in this [in TOF], otherwise I wouldn't have come because there is nothing to suggest that I need to see a doctor. So when I left I thought it was a bit strange. It was pointless for me to go there when there was nothing to come for. (Citizen 2).

In parallel with these reservations regarding the stratification model other GPs and municipal health professionals reported positive experiences. A challenge in assessing the stratification model was that the citizens attracted to the pilot implementation were primarily those who already had sufficient health competences. Thus, the citizens seen by the GPs might not be a representative subset of the citizens stratified to be at high risk, thereby biasing the GPs' assessment of the stratification model. While the GPs considered these citizens erroneously stratified, some of the citizens with good health competences saw the appointment with their GP as a reassuring component in taking care of their own health:

I have just been reassured that we eat okay and get the exercise we need. That, however, was also what I needed to know. I like going to spin class and getting my heart rate up, and there are those who say that it is not healthy to exert yourself like that. But, I was told that it is fine to do so, even though I am close to sixty. (Citizen 3).

In contrast, the GPs suspected that the citizens with low health competences, very unhealthy behaviors, and psychological conditions (e.g., depression) tended to opt out. Failing to attract these citizens was problematic because they would benefit from a healthier lifestyle and because the GPs and municipal health professionals felt that they had something to offer these citizens. As it were, they were let down:

I think we let the psychiatric citizens down because they don't come to us, and many of them take medication that makes them obese. We sometimes get them in to take blood tests because of their medication, and sometimes we also do ECGs because the medication may affect their heart rhythm. But otherwise we don't see them, and many have an unhealthy lifestyle. We don't find those with diabetes. (GP 1).

The discontent with the stratification model and the misfit between the attracted citizens and the current health offers were learning outcomes that challenged central elements of TOF. By making these challenges apparent prior to large-scale implementation, the pilot implementation showed its value and provided an opportunity for adjusting the TOF system. Both challenges had to be addressed in the recoupling of TOF with the embedding healthcare infrastructure. First, the stratification model had to be adjusted to align better with the preventive health offers provided by the GPs and municipal health centers. This adjustment might also involve adapting the role of the GPs and municipal health professionals in preventive care. Second, the outreach to citizens with low health competences had to be improved. To address this challenge new knowledge was needed.

4.8. Maintaining Momentum

During the spring of 2017, the TOF steering committee scrutinized the learning outcome of the pilot implementation. An immediate idea for addressing the discontent with the stratification model was to adjust the thresholds that determined when a citizen was stratified to be at high and moderate risk. The thresholds were, however, set on the basis of scientific evidence; thus, adjusting them might reduce the clinical precision of the stratification. Gradually, the steering group came to the conclusion that the problem was not the stratification model but the way it had been applied in defining the intervention. This conclusion shifted the focus toward adjusting the organizational model, for example by making the municipal health centers responsible for more of the intervention and by defining new responsibilities for the nurses in the GP clinics. It was specifically considered to task the nurses with conducting the majority of the health interviews, thereby restricting the GPs' involvement to the severe cases.

Input to the adjustments of the TOF system was sought through workshops. For example, a series of workshops were conducted to link the health folder closer to the needs and prerequisites of its users: citizens, municipal health professionals, and personnel from the GP clinics. These workshops resulted, among other things, in the removal of medical terms and formulations from the parts of the health folder directed at the citizens. The workshops also served to facilitate the participants in articulating how the TOF system could be better integrated in their daily practices. While the workshops provided useful insights about the organizational and personal context of the TOF system, they were by their very nature restricted in scope, duration, and realism. The original intention was for the pilot implementation to point out issues that would quickly be resolved, and then the project would be scaled up to more municipalities for a longer period of time. However, the analysis of the learning from the pilot implementation revealed that the required adjustments were more substantial than expected. As a result, the municipalities announced that without a deeper understanding of the reasons for the problems encountered during the pilot implementation they were disinclined to scale up the project. It was finally decided that the next step in TOF should be a second pilot implementation (in the fall of 2018).

The decision to conduct a second pilot implementation revealed some of the work required to enact the TOF system. First, the second pilot implementation was necessary as a means of maintaining the legitimacy of the project with its core stakeholders, who requested a better fit with current practices. Second, it showed the necessity of maintaining TOF as a temporary organization to make more time for recoupling it with the embedding healthcare infrastructure. Third, it maintained momentum in a situation where the original intention of a quick recoupling followed by large-scale implementation proved infeasible.

5. Discussion

Pilot implementations provide possibilities for bootstrapping new systems. These possibilities are shaped by temporal issues and central to what can be gained and otherwise learned from pilot implementations. In the following we discuss the learning that happens in pilot implementations, the bootstrapping of healthcare systems, and the implications of this study.

5.1. Learning in Pilot Implementations

Previous research finds that it is often difficult to learn from pilot implementations (Hertzum et al., 2019), mainly because the pilot implementation does not become sufficiently decoupled from the embedding organization to maintain a focus on learning. It is therefore worth noting that the TOF pilot implementation was successful in the sense that it produced important learning about the TOF system. This finding accords with previous studies finding that temporary isolation creates room for learning (Bakker et al., 2011; Scarbrough et al., 2004; Zietsma and Lawrence, 2010). However, the learning from the TOF pilot implementation did not lead to large-scale implementation but to the extension of the temporary state with a second pilot implementation to learn more.

The temporariness of the pilot implementation constrained the consequences of failure, in case the stratification model or another TOF component turned out to be dysfunctional. Thus, the pilot implementation could be allowed the freedom to employ yet unproven solutions. However, the temporariness of the TOF pilot implementation also constrained it in two unfavorable ways. First, any large-scale implementation of the TOF system presupposes a model for reimbursing the GPs, but the financial basis of the pilot implementation was a local agreement. This explicitly temporary agreement exploited that the pilot implementation was defined as bracketed from the embedding healthcare infrastructure. A local agreement was partly the normal way of financing GP participation in innovation projects and partly the best possible solution given the poor negotiation climate between RSD and OGP at the time. In this way the pilot implementation escaped the inertia of the embedding healthcare infrastructure but left a lasting financial model for the TOF system unresolved. Second, the three-month duration of the period of pilot use made it impossible to adjust the TOF system midway to fix problems they had already learned about. For example, it would have required a renegotiation of the local agreement to shift some of the health interviews from the GPs to the general-practice nurses. While the need for restricting their involvement was perceived by the GPs well before the end of the three-month period, it was not an option to renegotiate the local agreement in such a short period of time. In this way the infrastructural complexity made the pilot implementation less agile, because any change involved wrestling with the installed base (Star and Ruhleder, 1996). The absence of a financial model and midway adjustments exemplifies that the TOF pilot implementation was in some ways bracketed from the embedding healthcare infrastructure and in other ways interconnected with it. The decoupling was neither absolute, nor inconsequential.

Compared to many pilot implementations of information systems the TOF pilot implementation was large and involved diverse users: two municipalities, 47 GPs, and 2661 citizens over a period of three months. The large size was motivated by the TOF stakeholders' inclination to evidence-based thinking. As a result, the learning was probably less messy than the learning reported from other pilot implementations (Hertzum et al., 2019), partly because the size of the TOF pilot implementation provided a solid basis for learning and partly because the stratification and organizational models were well-defined. Thereby, TOF avoided some of the difficulty in telling the particulars of the pilot implementation from the lasting qualities of the system. The ability to make this distinction differed however across user groups, who as a consequence perceived the opportunities for learning and change quite differently. In particular, the citizens perceived TOF as a fixed entity that they either appreciated or found somewhat confusing and pointless, whereas the GPs perceived TOF as a malleable entity that could be changed in response to their feedback. This finding resembles Winthereik's (2010) study of the pilot implementation of an electronic maternity care record.

While part of the learning resulted from the three months of pilot use, another part resulted from the three years of preparations that set the TOF pilot implementation in motion. It is well-known that learning also happens during the preparatory stages of pilot implementations (Hertzum et al., 2012, 2019) but we find that this learning was of a different kind than the learning that resulted from the period of pilot use. The learning from the preparations was about gaining legitimacy and obtaining alignment. That is, it involved the active construction of a shared understanding or negotiated agreement among the stakeholders in TOF. The stakeholders needed to engage in alignment work (Jackson et al., 2011) to conceive, discuss, examine, mesh, revise, and settle on a mutually acceptable design of the TOF system, at least temporarily. Because this learning concerns efforts to create the conditions for preventive care to evolve, it assigns pilot implementations a generative role. In contrast the learning during the period of pilot use was about validating the precision and feasibility of the stratification and organizational models. This kind of learning tends to be focal in previous studies of pilot implementation (e.g., Hertzum et al., 2012; Pal et al., 2008; Rzevski, 1984); it assigns pilot implementations the more restricted role of a test.

The validation of the stratification and organizational models was an important learning outcome of the pilot implementation but we contend that the generative learning was more decisive. Specifically, the pilot implementation showed that TOF was positioned in between powerful stakeholders with diverging opinions about the evolution of preventive care. It was no small feat to maneuver these stakeholders into alignment and, thereby, enact the TOF system. One of several complications was the Sentinel case, which created a messy period in the preparation of the pilot implementation. It was key to the pilot implementation that the generative learning not just emerged during the preparations but also was applied during them. For example, the organizational model was adjusted by making it voluntary for the GPs to participate. This change was introduced after it was learned that the focus of TOF on lifestyle-related disease was met with resistance from some GPs, who were concerned about the risk of adverse effects of preventive health offers (pathologization). In contrast, the validation learning was not applied until after the pilot implementation. The application of generative learning during pilot implementations positions them as a potentially important method for bootstrapping systems.

5.2. Bootstrapping Systems

A prerequisite for infrastructure evolution is that extensions manifest managerial alignment and attract a critical mass of users (Hanseth and Aanestad, 2003; Hanseth and Lyytinen, 2010). This requires persistent efforts from designers because the process rarely proceeds as planned. For example, Aanestad and Hanseth (2002) highlight that projects make detours from the grand plan to form alliances, perform stunts to showcase the project, and reap spillover effects of these moves through the accumulation of knowledge. The purpose of such moves is to bootstrap the extension of the infrastructure. In TOF the long period that preceded the period of pilot use involved several stunts, including the feasibility study to prove the clinical precision of the stratification model. TOF also took several detours, including the replacement of Sentinel to accommodate OGP's skepticism toward it. While these deviations from the original plan were to a large extent responses to contingent events, they also produced spillover knowledge about how to navigate the politically textured field of preventive care. In particular, it improved the TOF project group's understanding of opportunities inherent in the embedding healthcare infrastructure. For example, it became clear that the installed base also encompassed opinions at the political level about how to manage the data extraction and that TOF could align with these opinions by developing a more transparent tool for extracting the data required to stratify the citizens.

The efforts to bootstrap the TOF system left a trace of events that, while they were responses to immediate constraints, strengthened the long-term alliance between the stakeholders. This alliance gave the project sufficient momentum and created a sense of irreversibility among the political stakeholders (Granqvist and Gustafsson, 2016) which enabled them to sustain the long, difficult periods and proceed towards the period of pilot use of the system. In preparing the system for pilot use it, however, became evident that it could not be restricted to a small user population, as RSD expected an evaluation based on a number of citizens, which in turn required the effort of several GPs and two municipal health centers to achieve. By introducing the system to a large network of users, TOF diverged from the typical recommendation of bootstrapping a new system, by focusing design on simple IT capabilities that offer immediate usefulness to a small user population in order to create an attractor that enable further, incremental scaling (Hanseth and Lyytinen, 2010; Hanseth and Aanestad, 2003). Yet, the feasibility study, the replacement of Sentinel, the sealed local agreement, the shift to voluntary GP participation, the enrolment of citizens, the period of pilot use, and the many other pilot-implementation activities collectively convinced the stakeholders that the project group was willing and able to adapt to the embedding infrastructure and that the TOF system moved forward, even if slowly. This way, the pilot implementation served to bootstrap the TOF system. The bootstrapping also benefited from the temporary bracketing of the pilot implementation from the embedding healthcare infrastructure. This planned isolation (Lundin and Söderholm, 1995) enabled real use of the system before RSD, OGP, or any other actor had to commit firmly to the system. None of the key actors would have been prepared to make such a commitment. Thus, the possibility to subject the TOF system to real use was predicated on the explicit temporariness of the use period: RSD, OGP, and the other actors merely had to commit to a process with a fixed start and end date. This reduced commitment enabled the pilot implementation, which in turn generated momentum and will help focus future efforts on the remaining obstacles.

While current literature on pilot implementations mainly highlights the importance of defining their scope, preparing the system and site, trialing the system, and learning from the experience (Hertzum et al., 2012), the bootstrapping of the TOF system required that the pilot implementation was bracketed from the embedding infrastructure. This was achieved by building a temporary organization, which included project staff, working groups, and dedicated funding from RSD. As argued by Lundin and Söderholm (1995) the recoupling of a temporary organization with the embedding permanent organization is crucial to the transfer of the learning that has been generated. The TOF pilot implementation shows that this recoupling is strongly dependent on obtaining alignment among the stakeholders during the preparations and period of pilot use, not just on collecting concrete evidence. As it were, the concrete evidence (e.g., the stratification of the citizens) did not satisfy the municipalities and GPs, thereby making the alliance among the stakeholders too fragile to proceed immediately to large-scale implementation. However, the alignment obtained during the pilot implementation gave the stakeholders a sense of cohesiveness and the impetus to continue their efforts. To create more time for bootstrapping, the objective became to prolong or postpone recoupling, rather than to enforce it. In this way recoupling was turned into an extended process. The temporal extension of the recoupling process was achieved by the decision to conduct a second pilot implementation, that is, by arguing the need for a deeper understanding of the organizational constraints for preventive care in general practice. This shows the double need for decoupling systems to bootstrap them through stunts and detours and for keeping them decoupled long enough for the bootstrapping to take place.

5.3. Implications

In summary, we see five implications of this study. First, pilot implementations provide a means for bootstrapping systems by working through all the steps necessary to try out the system in the pilot setting. This way a pilot implementation may facilitate the evolution of the embedding infrastructure, or reveal that the system is not ready for incorporation in the infrastructure. To achieve these valuable ends pilot implementations must be clearly separated from the first stage in an incremental, full-scale implementation. We contend that this requires more research on pilot implementations, more managerial awareness of their qualities, and more use of them in practice.

Second, the preparatory stages of a pilot implementation may be more about gaining legitimacy and obtaining alignment than about technical configuration and data migration. Activities such as technical configuration and data migration suggest that the preparations are a phase that ends before the period of pilot use can begin. That is, they are consistent with an understanding of pilot implementations as temporary organizations, which serialize time into before, during, and after. In contrast, legitimacy and alignment suggest the presence of multiple parallel timelines (Steinhardt and Jackson, 2014) that can be influenced by the pilot implementation but also exist independently of it. That is, legitimacy and alignment emphasize infrastructure time over project time (Karasti el at., 2010). Practitioners are continually faced with the challenge of reconciling these two temporal perspectives and should utilize pilot implementations to gain the legitimacy and alignment necessary to bootstrap systems.

Third, the temporariness and restricted scale of pilot implementations yield degrees of freedom that improve the conditions for innovation and learning. At the same time, solutions to key challenges may be tied to the temporariness and thereby not valid beyond the pilot implementation. Future research should elaborate the mechanisms through which pilot implementations facilitate infrastructure evolution, including the role of their explicit temporariness in fostering long-term change. Existing studies of infrastructure evolution tend to emphasize its extended timescale and embeddedness in ongoing use (e.g., Karasti et al., 2010; Ribes and Finholt, 2009), thereby deemphasizing explicitly temporary efforts.

Fourth, the boundaries of pilot implementations are continuously created and recreated through decoupling and recoupling. Research on temporary organizations (Burke et al., 2016) should embrace that decoupling and recoupling are constant sources of friction, not merely phases at the beginning and end of a temporary organization. Infrastructure research contributes an attention to the reach, scale, and interwovenness of the numerous relations that must be decoupled and recoupled during a pilot implementation. It appears particularly important to understand better the continuous character of recoupling because it may offer insights into the difficulties that often ensue in bringing the learning from a temporary organization home to the embedding organization (Bakker et al., 2011).

Fifth, pilot implementations have a generative role in infrastructure evolution; they are not merely tests. This suggests hitherto unexplored relations between pilot implementations and methods such as living labs (e.g., Ley et al., 2015). Exploring these relations might, for example, provide insights about how to make pilot implementations more agile during the period of pilot use.

5.4. Limitations

Three limitations should be remembered in interpreting the results of this study. First, the TOF pilot implementation is inextricably woven into the Danish healthcare infrastructure with its powerful actors at multiple levels, complex public reimbursement of the GPs, and shortage of GPs in many municipalities. We acknowledge the need for future research to investigate the generative role of pilot implementations in non-Danish and non-healthcare settings. Second, the study is not an assessment of the entire TOF project. We investigate the pilot implementation, including the yearlong preparations leading up to the period of pilot use. The project continues after the pilot implementation and we, thus, cannot know what will ultimately be learned from it. Third, the citizens' perspective on preventive care remains a supplementary topic. In particular, we do not investigate their views on how preventive care offers can better reach and include citizens with low health competences. These citizens tended to opt out of the TOF pilot implementation, though they were central to its target group.

6. Conclusion

The temporariness of pilot implementations is central to their role in infrastructure evolution because it enables experimentation and learning. In the TOF pilot implementation the learning was about creating the conditions for the TOF system as well as about validating it. While the generative learning led to legitimacy and alignment, the validation learning yielded clinical evidence and first-hand experience. The temporariness also allowed for temporary solutions to issues that otherwise blocked progress, thereby exporting the issues to the recoupling process but also exporting a temporary alignment that will facilitate the stakeholders in reaching long-term solutions. In this way pilot implementations enable the enactment of systems by bootstrapping them. Pilot implementations are conducive to bootstrapping because they, if successful, render alignment manifest, make benefits salient, and create the momentum necessary to start off the system. To succeed, the decoupling and recoupling of the temporary organization must be approached as constant sources of friction; they are not merely the beginning and end of the pilot implementation. While the decoupling is anything but absolute, the pilot implementation must remain decoupled long enough for bootstrapping to take effect. We conclude that a pilot implementation can contribute to the bootstrapping of a system by providing a temporary organizational framing for enacting it.

Acknowledgements

This paper is a revised and thoroughly extended version of a paper presented at the InfraHealth2017 workshop on infrastructures for healthcare. In the interest of full disclosure, we declare that the third author has been the head of RUGP during the TOF project and involved in managing the project. We are grateful to [names blinded for review] for support in identifying interviewees with a stake in the pilot implementation. Special thanks are due to the interviewees.

References

Aanestad, Margunn; and Ole Hanseth (2002). Growing networks: Detours, stunts and spillovers. In *COOP 2002: Proceedings of the Fifth International Conference on Designing Cooperative Systems, 4–6 June 2002, St. Raphael, France*, pp. 38–48.

Aanestad, Margunn; and Tina B. Jensen (2011). Building nation-wide information infrastructures in healthcare through modular implementation strategies. *Journal of Strategic Information Systems*, vol. 20, no. 2, pp. 161–176.

Aanestad, Margunn; Miria Grisot; Ole Hanseth; and Polyxeni Vassilakopoulou (2017). Information infrastructures and the challenge of the installed base. In M. Aanestad, M. Grisot, O. Hanseth and P. Vassilakopoulou (eds), *Information infrastructures within European health care: Working with the installed base*. Cham, Switzerland: Springer, pp. 25–33.

Bakker, Rene M.; Bart Cambré; Leonique Korlaar; and Joerg Raab (2011). Managing the project learning paradox: A set-theoretic approach toward project knowledge transfer. *International Journal of Project Management*, vol. 29, no. 5, pp. 494–503.

Bakker, Rene M.; Robert J. DeFillippi; Andreas Schwab; and Jörg Sydow (2016). Temporary organizing: Promises, processes, problems. *Organization Studies*, vol. 37, no. 12, pp. 1703–1719.

Bowker, Geoffrey C.; and Susan L. Star (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Brown, John S.; and Paul Duguid (2001). Knowledge and organization: A social-practice perspective. *Organization Science*, vol. 12, no. 2, pp. 198–213.

Burke, Catriona M.; and Michael J. Morley (2016). On temporary organizations: A review, synthesis and research agenda. *Human Relations*, vol. 69, no. 6, pp. 1235–1258.

Buxton, M. J. (1987). Problems in the economic appraisal of new health technology: The evaluation of heart transplants in the UK. In M. F. Drummond (ed.): *Economic appraisal of health technology in the European Community*. Oxford: Oxford Medical Publications, pp. 103–118.

Chen Nan-Chen; Sarah S. Poon; Lavanya Ramakrishnan; and Cecilia R. Aragon (2016). Considering time in designing large-scale systems for scientific computing. In *Proceedings of the CSCW2016 conference on computer supported cooperative work and social computing*. New York: ACM Press, pp. 1535–1547.

Glasgow, Russell E.; C. Tracy Orleans; Edward H. Wagner; Susan J. Curry; and Leif I. Solberg (2001). Does the chronic care model serve also as a template for improving prevention? *Milbank Quarterly*, vol. 79, no. 4, pp. 579–612.

Grabher, Gernot (2002). Cool projects, boring institutions: Temporary collaboration in social context. *Regional Studies*, vol. 36, no. 3, pp. 205–214.

Granqvist, Nina; and Robin Gustafsson (2016). Temporal institutional work. *Academy of Management Journal*, vol. 59, no. 3, pp. 1009–1035.

Hanseth, Ole; and Margunn Aanestad (2003). Design as bootstrapping: On the evolution of ICT networks in health care. *Methods of Information in Medicine*, vol 42, no. 4, pp. 385–391.

Hanseth, Ole; and Kalle Lyytinen (2010). Design theory for dynamic complexity in information infrastructures: The case of building internet. *Journal of Information Technology*, vol. 25, no. 1, pp. 1–19.

Hertzum, Morten; Jørgen P. Bansler; Erling Havn; and Jesper Simonsen (2012). Pilot implementation: Learning from field tests in IS development. *Communications of the Association for Information Systems*, vol. 30, no. 1, pp. 313–328.

Hertzum, Morten; Maria I. Manikas; and Arnvør Torkilsheyggi (2019). Grappling with the future: The messiness of pilot implementation in information systems design. *Health Informatics Journal*, vol. 25, no. 2, pp. 372–388.

Jackson, Steven J.; David Ribes; Ayse Buyuktur; and Geoffrey C Bowker (2011). Collaborative rhythm: Temporal dissonance in collaborative scientific work. In *Proceedings of the CSCW2011 conference on computer supported cooperative work*. New York: ACM Press, pp. 245–254.

Karasti, Helena; Karen S. Baker; and Florence Millerand (2010). Infrastructure time: Long-term matters in collaborative development. *Computer Supported Cooperative Work*, vol. 19, no. 3–4, pp. 377–415.

Ley, Benedikt; Corinna Ogonowski; Mu Mu; Jan Hess; Nicholas Race; David Randall; Mark Rouncefield; and Volker Wulf (2015). At home with users: A comparative view of living labs. *Interacting with Computers*, vol. 27, no. 1, pp. 21–35.

Ligthart, Rik; Leon Oerlemans; and Niels Noorderhaven (2016). In the shadows of time: A case study of flexibility behaviors in an interorganizational project. *Organization Studies*, vol. 37, no. 12, pp. 1721–1743.

Lundin, Rolf A.; and Anders Söderholm (1995). A theory of the temporary organization. *Scandinavian Journal of Management*, vol. 11, no. 4, pp. 437–455.

Orlikowski, Wanda J.; and JoAnne Yates (2002). It's about time: Temporal structuring in organizations. *Organization Science*, vol. 13, no. 6, pp. 684–700.

Pal, Raktim; Arijit Sengupta; and Indranil Bose (2008). Role of pilot study in assessing viability of new technology projects: The case of RFID in parking operations. *Communications of the Association for Information Systems*, vol. 23, article 15, pp. 257–276.

Pipek, Volkmar; and Volker Wulf (2009). Infrastructuring: Toward an integrated perspective on the design and use of information technology. *Journal of the Association for Information Systems*, vol. 10, no. 5, pp. 447–473.

Randell, David P.; E. Ilana Diamant; and Charlotte P. Lee (2015). Creating sustainable cyberinfrastructures. In *Proceedings of the CHI2015 conference on human factors in computing systems*. New York: ACM Press, pp. 1759–1768.

Ribes, David; and Thomas A. Finholt (2009). The long now of technology infrastructure: Articulating tensions in development. *Journal of the Association of Information Systems*, vol. 10, no. 5, pp. 375–398.

Rzevski, G. (1984). Prototypes versus pilot systems: Strategies for evolutionary information system development. In R. Budde, K. Kuhlenkamp, L. Mathiassen, and L. Zullighoven (eds.): *Approaches to prototyping: Proceedings on the working conference on prototyping*, Heidelberg, Springer, pp. 356–367.

Scarbrough, Harry; Jacky Swan; Stéphane Laurent; Mike Bresnen; Linda Edelman; and Sue Newell (2004). Project-based learning and the role of learning boundaries. *Organization Studies*, vol. 25, no. 9, pp. 1579–1600.

Schatz, Bruce R.; and Richard B. Berlin (2011). Healthcare infrastructure: Health systems for individuals and populations. London: Springer.

Star, Susan L.; and Karen Ruhleder (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research*, vol. 7, no. 1, pp. 111–134.

Steinhardt, Stephanie B.; and Steven J. Jackson (2014). Reconciling rhythms: Plans and temporal alignment in collaborative scientific work. In *Proceedings of the CSCW2014 conference on Computer Supported Cooperative Work & Social Computing*. New York: ACM Press, pp. 134–145.

Steinhardt, Stephanie B.; and Steven J. Jackson (2015). Anticipation work: Cultivating vision in collective practice. In *Proceedings of the CSCW2015 conference on computer supported cooperative work and social computing*. New York: ACM Press, pp. 443–453.

Stjerne, Iben S.; and Silviya Svejenova (2016). Connecting temporary and permanent organizing: Tensions and boundary work in sequential film projects. *Organization Studies*, vol. 37, no. 12, pp. 1771–1792.

Strauss, Anselm; and Juliet Corbin (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory. Second Edition. Thousand Oaks: Sage.

Suchman, Lucy A. (2007). Human-machine reconfigurations: Plans and situated action, 2nd edition. Cambridge, UK: Cambridge University Press.

Winthereik Brit R. (2010). The project multiple: Enactments of systems development. *Scandinavian Journal of Information Systems*, vol. 22, no. 2, pp. 49–64.

Zerubavel, Eviatar (1979). Patterns of time in hospital life: A sociological perspective. Chicago, Ill: University of Chicago Press.

Zietsma, Charlene; and Thomas B. Lawrence (2010). Institutional work in the transformation of an organizational field: The interplay of boundary work and practice work. *Administrative Science Quarterly*, vol. 55, no. 2, pp. 189–221.