

Designing Measures for Managing the Effectiveness for Integrated Service Delivery: The Case of a Geriatric Outpatient Clinic

Virpi Sillanpää, Tampere University of Technology, Tampere, Finland

ABSTRACT

Health and social services are increasingly produced in cooperation among several specialized organizations. This has increased the importance of service integration in the sector. While the literature acknowledges the need for performance measurement addressing outcomes and effectiveness of service integration, not enough is known about applying such measurement practices. This paper examines the design of a performance measurement system to support the effectiveness management of an integrated service delivery. The research identifies three aspects of effectiveness – community, client and network level. Empirical examination reveals that success factors for design of measures for integrated service delivery include consensus on the targets, coordination and fluent cooperation among actors in the network. The paper contributes to performance measurement literature by illustrating how the design of system level measurement is carried out in practice and analyzing the lessons learned.

Keywords: Case Study, Effectiveness, Finland, Performance Management, Performance Measurement, Service Integration, Welfare Services

INTRODUCTION

Need for System-level Management of Welfare Services

Effectiveness has become a focal performance criterion in health and social services. One reason for this has been the New Public Management, which emphasizes the need to modernize the public sector. Public organizations are expected to be managed more like enterprises and become more customer-oriented, focused on outcomes rather than inputs, and be more efficient and effective (Barretta & Busco, 2011; Jansen, 2008).

Outcomes and effectiveness in many fields of welfare are produced in cooperation among different organizations in the health and social sectors (e.g. Axelsson & Axelsson, 2006; Hansson et al., 2010). Since the 1990's increased attention has been paid to cooperation as a means to achieve more effective and efficient public sector services (e.g. Kurunmäki & Miller, 2011). Along with the increasing cooperation, service integration has become an important issue in welfare services

DOI: 10.4018/IJPPHME.2015010103

(e.g. Axelsson & Axelsson, 2006; Hansson et al., 2010). Integrated service delivery strategies are deemed a solution to rising costs, low quality of care and dissatisfied clients (Evans & Baker, 2012), and in recent years much research on collaboration (e.g. Qvretveit, 2002) and the evaluation of integration of services has been conducted (e.g. Ahlgren & Axelsson, 2005; Axelsson & Axelsson, 2006; Standberg-Larsen & Krasnik, 2009). The management and coordination of intra- and inter-organizational cooperation are challenging, and research generally aims to find concrete tools and guidelines for collaboration and service integration (e.g. Qvretveit, 2002; Hansson et al., 2010). Better tools for managing integration are needed, since many integration efforts have failed (Qvretveit, 2002). There may be concern that service integration has become means without ends or an end in itself. The academic literature emphasizes process rather than outcome in initiatives to apply service integration in practice, ultimately producing relatively scant evidence of the outcomes of integration (Wistow & Dickinson, 2012).

Increasing cooperation and integration in health care and other welfare services indicate that performance management at the organizational level is not sufficient to ensure high performance of the service system. Organizations may have conflicting goals and focus primarily on their own performance rather than the overall ability of the system to serve customers. This puts performance management activities in a new perspective. Managers need information on system-level performance (e.g. Callender, 2011). This has been recognized in public management literature and the focus is moving towards inter-organizational governance, which emphasizes the role of long-term inter-organizational relationships and the governance of processes as solutions for more effective public services (Osborne, 2006). However, limited attention has been paid to studying how management control practices like performance measurement are applied in practice within inter-organizational relationships, or what public network performance is (Barretta & Busco, 2011).

Effectiveness of welfare services at the service system level is usually assessed by utilizing different types of evaluation approaches. The role of economic evaluation is established especially in the field of healthcare (Drummond et al., 2005; Brazier et al., 2007), but applied also in social services (Flatau and Zaretsky, 2008; Jones et al., 1994). Key motivation for conducting economic evaluation is its ability to provide systematic analysis of different alternatives in service production for decision makers at the service system level (i.e. public administration). Economic analyses seek to identify and to make explicit the set of criteria that is useful in deciding among different uses of scarce resources. Key characteristics of economic evaluations, regardless the area it is applied in features the inclusion of the costs and consequences of activities (Drummond et al., 2005, p. 9). Focal benefits of evaluation approach relate to its ability to provide information of effectiveness of operation. However, since evaluations are usually conducted on an ad hoc, retrospective basis by external evaluators, and results are communicated in the format of extensive evaluation reports, their usability as managerial tools is limited (Nielsen & Ejler, 2008). Managerial purposes require different types of performance measurement systems. Performance measurement, which can be defined as *“the process of defining, monitoring, and using objective indicators of the performance of organizations and programs on a regular basis”* (Poister, 2003, p. 1) tracks the results as an integral part of operation and is conducted by internal managers, appears as a more applicable tool for managing effectiveness (Nielsen & Ejler, 2008). Next, performance measurement as theoretical approach of this research is discussed.

Theoretical Approach of the Study and Research Gap

Performance measurement is the crucial element of performance management, which can be defined as a process by which the organization manages its performance in line with its strategies and objectives (Bititci et al., 1997). As a diverse and multidisciplinary research area, the

field of performance measurement lacks a cohesive body of knowledge (e.g. Franco-Santos et al., 2007). Thus, in the field of performance measurement, theoretical approach refers to contributions made by management researchers in various disciplines, like in strategy management, operations management and management accounting rather than solid knowledge within a certain discipline (ibid). While performance measurement appears as a promising approach in managing effectiveness, there are several gaps concerning the current knowledge of performance measurement literature related to measuring the effectiveness of integrated service delivery and welfare service systems (e.g. Barretta & Busco, 2011; Conaty, 2012; Provan & Milward, 2001). Traditionally performance measurement research is focused on organizations. Recently the performance measurement of inter-firm cooperation and networks have received increasing attention both in business (e.g. Busi & Bititci, 2006; Kaplan et al., 2010; Parung & Bititci, 2008; Pekkola, 2013) and the public sector literature (e.g. Murray et al., 2000; Conaty, 2012). Simultaneously, the performance focus in the research literature has shifted from improving processes and outputs to longer-term outcomes and impacts and the value that co-operative service delivery creates for its clients (e.g. Barretta & Busco, 2011; Halligan et al., 2012; Laihonen et al., 2014; Porter 2010). In spite of recent development in the area of performance measurement research, there is very little empirical research that explores the performance measurement and management in collaboration (e.g. Bititci et al., 2012), especially in the area of welfare services.

Earlier research on how to carry out a successful measurement project at organizational level abound (e.g. Jääskeläinen, 2010; Bourne et al., 2000; Neely et al., 2005; Rantanen et al., 2007; Saunila et al., 2012). However, little is currently understood about how to develop performance measurement that captures effectiveness in inter-organizational co-operative settings in the welfare sector. Examples from business (e.g. Kulmala & Lönnqvist, 2006; Pekkola, 2013) are not necessarily applicable to public sector welfare services for many reasons; for example in the public sector partnerships and networks are usually imposed by superior authority rather than resulting from voluntary cooperation, which may affect the motivation and interest of participants for development work (Halligan et al. 2012). Common challenges in social and healthcare service networks include unclear roles and responsibilities (e.g. Hansson et al., 2010), which challenges the start of performance measurement in inter-organizational cooperation. In public service networks with various stakeholders and levels of analysis even the content of performance appears unclear (e.g. Barretta & Busco, 2011; Conaty, 2012). In efforts to measure the effectiveness of integrated service delivery, the number of stakeholders with varying targets and performance criteria makes even agreeing on a holistic and balanced definition of effectiveness difficult (e.g. Conaty, 2012; Evans & Baker, 2012; Provan & Milward, 2001).

To conclude, there appears to be lack of knowledge on applying performance measurement approach in measuring effectiveness of welfare services at system level. This research aims at fulfilling the research gap by examining how the performance measurement approach can be applied to measure the effectiveness of integrated service delivery in welfare services. This research focuses on the *design phase of measurement*, since the first steps of measurement are the most critical for the successful use of performance measurement later on (e.g. Bourne et al., 2000; Rantanen et al., 2007). Performance measurement design entails 1) consideration of what to measure and then 2) how to carry out the measurement development (e.g. Bourne et al., 2000; Jääskeläinen et al., 2009). In this research, defining *elements of effectiveness* of the service system is a prerequisite for specifying the aspects that should be measured. Another focal question is to understand how the measurement design process can be implemented in practice in the context of inter-organizational cooperation in the welfare service sector. This research contributes to the conceptual understanding on performance measurement design in the case of effectiveness in co-operative settings in welfare services. From a practitioner perspective, this

paper explains how to design measures of effectiveness that are relevant managerial tools at welfare service system level.

Next, the design process of performance measurement focusing effectiveness in welfare services is discussed based on earlier research literature.

LITERATURE REVIEW: PERFORMANCE MEASUREMENT DESIGN AT SERVICE SYSTEM LEVEL

Performance measurement implementation is usually divided into three main phases: design, implementation and use of measures (e.g. Bourne et al., 2000). The first task in designing measures is to identify the purpose of the measurement, which should be linked to the organization's objectives. In the context of networks, the measurement should support the objectives of the network (e.g. Bourne et al., 2000; Neely et al., 2005). Thus, the measurement design requires consensus on the objectives and the purpose of the measurement among actors in the network. (Cunha et al., 2008; Kaplan et al., 2010). Cunha et al. (2008) point out that defining indicators should be a joint effort during the network set-up.

In efforts to measure effectiveness, the consensus of the target of measurement, i.e. on the concept of effectiveness should be reached among network participants. Effectiveness is a challenging concept, especially regarding welfare services, since there are various key stakeholders ranging from individual clients (beneficiaries) to employees of service provider organizations and taxpayers, whose frequently conflicting needs and goals the effective service system should fulfill (Porter, 2010; Provan & Milward, 2001). Consistent with the multiple-stakeholder perspective, welfare service systems can be measured at different levels of analysis (e.g. Provan & Milward 2001; Crook et al., 2009; Strandberg-Larsen & Krasnik, 2009). Usually the frameworks proposed in the literature include three levels of analysis, for example, Provan & Milward (2001) distinguish levels of community, network and its organizational participants, whereas the classifications by Strandberg-Larsen & Krasnik (2009) and Crook et al. (2005) include system level, organization level, and individual level.

At the community level, the aim of the network usually relates to improved access, utilization, responsiveness and integration of services while maintaining or reducing costs. Thus, effectiveness at community-level can be evaluated by assessing the aggregate outcomes for clients of the network, and by estimating the overall costs of service (Crook et al., 2005; Provan & Milward, 2001). According to Crook et al. (2005), system-level outcomes include cost-savings across systems, lowering access barriers, networking among community organizations, and aggregation of client-level outcomes. The literature on welfare services suggests that aspects related both to client outcomes (e.g. changes in the quality of life, functional status, mortality) and costs have central role when measuring the overall performance of the system (Porter, 2010; Strandberg-Larsen & Krasnik, 2009). In addition to these, there are other criteria according to which community-level effectiveness can be measured, like the ability to build social capital and a public perception that a problem is being solved (Provan & Milward, 2001).

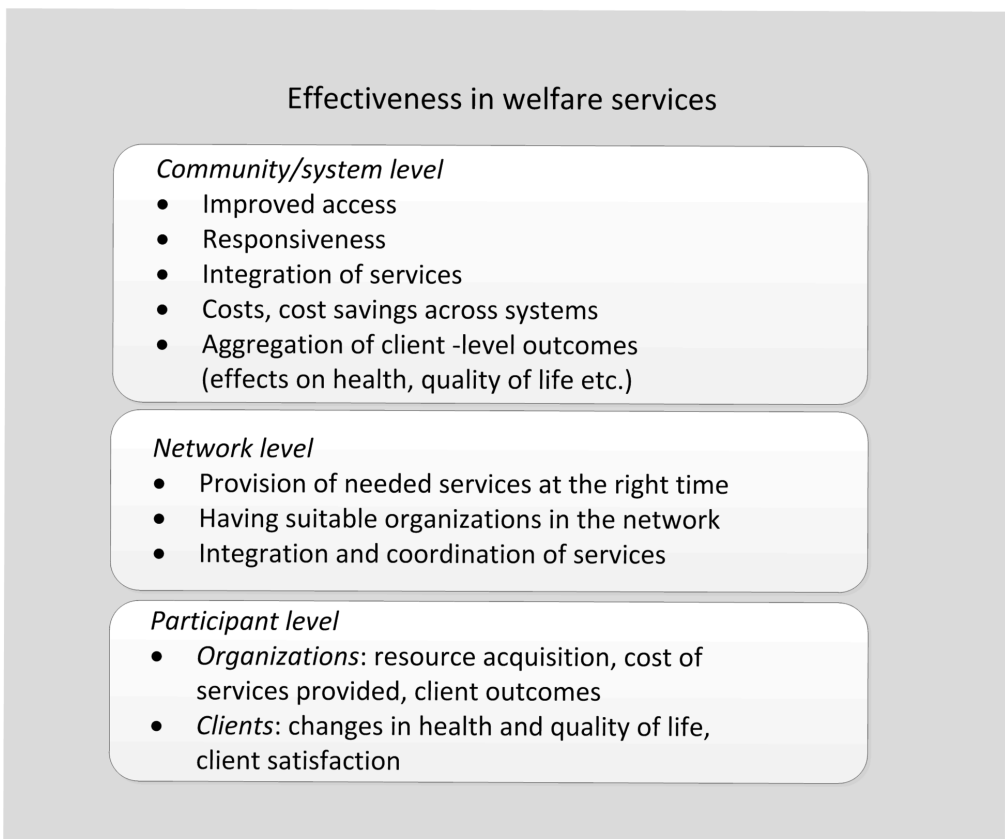
Effectiveness at the network level relates to the network's ability to provide its clients with the services needed at the right time. Effectiveness criteria at the network level include having suitable organizations providing essential services in the network, integration and coordination of services in order to avoid overlap and sub-optimization, creation and maintenance of actors coordinating the network, member commitment to network goals and cost of network maintenance (Provan & Milward, 2001). Effectiveness criteria at participant level include both criteria related to specific organizations and clients. Measures related to organizations in the network relate to

resource acquisition, cost of services provided, and client outcomes. The client-level includes measures related to changes in health and the quality of life of clients receiving services, and client satisfaction (e.g. Crook et al., 2005; Provan & Milward, 2001).

Based on the literature, Figure 1 proposes a conceptual framework for effectiveness in welfare services. Consonant with earlier literature (e.g. Crook et al., 2009; Provan & Milward 2001; Strandberg-Larsen & Krasnik, 2009) Figure 1 divides effectiveness into three levels of analysis: community/system level, network level, and participant level, which includes both organizations participating in the network and individual clients. The overall effectiveness of the welfare service system is based on interactions across all three levels of analysis (e.g. Crook et al., 2009), indicating that all three levels of analysis should be measured for a holistic view of the effectiveness of the system.

Usually the actual design of performance measurement is organized as a series of workshops led by a facilitator or consultant (Bourne et al., 2002; Lönnqvist, 2004; Pekkola, 2013). The facilitator's contribution is considered crucial to the success of the design process (Bourne et al., 2002; Jääskeläinen et al., 2013). In inter-organizational cooperation, the role of facilitator is further emphasized (e.g. Axelsson & Axelsson, 2006; Pekkola, 2013). The literature suggests that the successful design of performance measurement system assumes the participation of employees in the operative level in the development work (Jääskeläinen et al., 2013; Rantanen

Figure 1. Effectiveness in welfare services



et al., 2007). The composition of the workgroup and a positive atmosphere are considered focal factors promoting the development of performance measurement systems in both public welfare services (Jääskeläinen et al., 2013) and collaborative business networks (Kaplan et al., 2010; Pekkola, 2013). Inter-organizational collaboration in public health is often organized in the form of multidisciplinary teams, a small group of people representing different professions who together across organizational boundaries provide services to a specific group of clients. The active involvement of representatives from multidisciplinary teams is vital to the successful development of performance measures in welfare service systems (Axelsson & Axelsson 2006).

The design phase is critical in the development of any performance measurement system since it forms the foundation for the implementation and actual use of measures. Problems in the design phase may entail more problems later on (e.g. Jääskeläinen et al., 2009; Rantanen et al., 2007). Busi & Bititci (2006) point out that major barriers to the successful implementation of a collaborative performance management system relate to difficulty in developing a collaborative culture and in developing appropriate performance measures. According to Rantanen et al. (2007), the main challenges in designing performance measurement systems in the public sector are due to the involvement of many stakeholders, causing difficulties in accommodating the conflicting needs of stakeholders and difficulties in target setting. Public sector welfare service organizations have numerous measures in use (Jääskeläinen, 2010; Linna et al., 2010; Martin & Kettner, 2010; Sillanpää, 2011) and selecting appropriate measures from those available or developing new measures is challenging. Moreover, in the welfare service context, financial measures dominate (Saunila et al., 2012) even though the most focal measurement objects in health and social services relate to qualitative, intangible issues (e.g. satisfaction and well-being of clients, quality of life) that are difficult to measure (Jääskeläinen et al., 2009; Martin & Kettner, 2010; Sillanpää, 2011) and not measurable with financial measures. Other challenges in performance measurement development in the public sector include lack of ownership of the project, opposition of personnel to performance measurement (cultural issues), and overlapping/competing projects that impede the progress of the development work (Jääskeläinen, 2010; Rantanen et al., 2007).

Summarizing the literature, recent developments in welfare services (cooperation, new public management) underline the need to acquire knowledge about performance, and especially the effectiveness of services provided in cooperation. From the performance measurement perspective, the literature currently reveals little of how managers perceive effectiveness at the service system level (what to measure) and how a performance measurement system supporting the management of effectiveness in co-operative settings could be designed (how to measure) in the context of welfare services. In the next section, the empirical research attempts to answer these questions.

EMPIRICAL EXAMINATION: CONTEXT AND METHODS OF RESEARCH

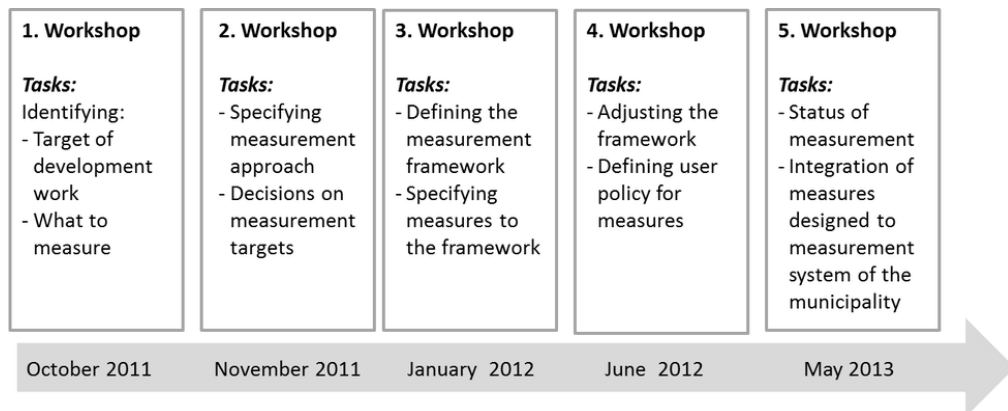
The empirical part of the research includes a qualitative case study (Gummesson, 2000) describing the development of a performance measurement system for a new type of service, the geriatric outpatient clinic Geripol, organized as inter-organizational cooperation in a large Finnish municipality. The practical aim of the development work was to design a performance measurement system to support the performance management of Geripol, especially in terms of quality and effectiveness. The municipality had a separate productivity measurement framework including costs and productivity. The aim of empirical research was to examine how the conceptual framework of effectiveness based on the research literature can be operationalized and if the identified factors related to design of performance measures, i.e. workshops as working

method, participants from multidisciplinary teams, role of facilitator, and challenges related to the design process, apply in the context of inter-organizational cooperation in welfare services. Case study was considered an appropriate research approach because the author sought a rich understanding about the real-life phenomena under study (Yin, 2009). The case study was conducted as an action research project with twofold objectives: solving a practical problem and contributing to prior knowledge (Coughlan & Coughlan, 2002). Action research, which can be seen as a variant of case research, is applicable in situations such as this, where the aim is to develop a new solution to a problem of which little is so far known. In action research, researchers act as participants “on site” rather than as outside observers, thereby gaining insights into unstructured situations and variables/issues that may not initially be apparent (Westbrook, 1995). The action research project was carried out in the form of workshops facilitated by two researchers. Research project included one group interview before the workshops and five workshops. Figure 2 illustrates the themes and schedules for workshops. This research approach afforded the researchers many-sided data and an opportunity to utilize many different types of sources of evidence, like administrative documents (reports, proposals), e-mail correspondence, interviews, and participant observation (e.g. Yin, 2009). The main data collection methods were participant observation during workshops and a group interview at the beginning of the project. In the workshops one of the researchers worked mainly as facilitator, whereas the other focused on collecting data by making notes on conversations during the workshop. After each workshop, the researchers analyzed the notes and made summaries of workshops that were discussed with participants at the beginning of the next workshop session. Thus participants had an opportunity to comment on the researchers’ observations and analyses, and data collection and analysis was an iterative process that continued throughout the development work. The researchers thereby accessed in-depth information about the development process and the concept of effectiveness and the participants’ interpretation of it. The author was one of the facilitators in the development project. The process and the participating actors are described in the next chapter.

The geriatric outpatient clinic was launched in 2011. The clinic’s services are produced through cooperation of various operational municipal service units, including the outpatient departments of three geriatric hospitals, primary health care, and home care, which are located in different administrative departments. Some services may be purchased from other service providers, like companies or non-profit organizations. Integrated service delivery, right from the referral of clients to supportive home care, including timely institutional care and specialized health care, is deemed a prerequisite for the successful operation of Geripol. One of Geripol’s main tasks is to coordinate service production so clients get the services they need at the right time. Clients’ service needs are assessed by Geripol, and clients are referred to the services they need, either produced by own units or purchased from companies. Coordinating the client referral process is one of the main changes that Geripol has entailed; client referrals were previously made by family doctors in primary health care. For the clients, coordination is the main benefit of Geripol as they now should receive all the services they need from one service counter.

The overall goal of Geripol is to enable elderly people to live at home and support home care, thereby reducing the use of institutional care. Home care is considered beneficial for both clients and service system; clients prefer living in their own homes and receiving the care and support they need there. The reduction in institutional care means substantial overall savings in elderly care. Besides the traditional outpatient clinic operations, Geripol aims to offer geriatricians’ services in clients’ homes, which would be conducive to living at home and also cater for clients in need of demanding specialized care. Housecalls made by a multidisciplinary team, including a geriatrician, a nurse and the client’s dedicated caregiver are the core of Geripol’s operations. One of the main advantages is the support the geriatrician offers the elderly person’s home care

Figure 2. Outline of workshops



team; consultation to the client's general practitioner in primary health care and training for home care staff. Dissemination of geriatric expertise and on-the-job learning are also among Geripol's targets. In addition to professional nursing staff, clients' family members are encouraged to be active (if possible) in supporting the client.

RESULTS

The Design Process in Geripol

The design of measurement was carried out as a development project comprising five workshops. Figure 2 illustrates the main tasks and the schedule of workshops, which were preceded by a group interview between the researchers and working group to survey the activities and targets of Geripol, current measurement practices and development needs regarding measurement. The working group in the municipality consisted of representatives of both the service provider and purchaser; the purchasing manager of elderly care services, the coordinator of elderly care services (purchaser), the manager of hospital care for elderly (provider), the development manager of institutional care (provider), the chief medical officer of geriatrics (consultant/developer of Geripol), and a professor of geriatrics serving as an expert in this project and also chairing another group developing efficiency measures for all elderly care services in the municipality. Workshops were organized by the research and evaluation manager of the municipality.

At the beginning of the development work, the task was to identify the target of the development project. A comprehensive picture of Geripol, its operations and stakeholders was drawn to illustrate the connections between the various organizations and stakeholders forming the network. The main target of the measurement development work was to design measures to evaluate the outcomes and effectiveness of Geripol services. Organizations in the network had numerous measures already in use for evaluating the effects of elderly care services, especially clinical measures assessing the health outcomes of institutional care. However, in addition to effects on clients, the working group identified several other targeted impacts related to Geripol, e.g. cost savings, intensified cooperation within the network, effects on institutional care, home care and primary health care, and changes in the competence of primary health care and home care personnel.

During workshops 2 - 4 the measurement approach was specified, which entailed identifying different perspectives related to the effectiveness of Geripol activities. During the development work, three main perspectives of effectiveness were identified: community, client, and management/production of services. Numerous measures were proposed and discussed in the workshops on the performance of Geripol. The aim of this development project was to create measures for the purchaser's use, which underlined the role of measures concerning effectiveness related to clients and community. According to the purchaser-provider model, the purchaser determines effectiveness, while the provider's task is to decide how to achieve it (how to manage the network). Hence only a few measures related to processes or cooperation were included in the framework. A definition of user policy for measures was proposed in the fourth workshop. In the user policy, the following issues were identified for each measure: name, measurement target (what to measure), formula for the measure, current value, target value, data needed for the measurement, person responsible for the measurement, and how the results were to be reported. In the final workshop the status of measurement and the integration of measures throughout the measurement system were discussed.

The duration of development project was relatively long (about 19 months), which influenced the composition of the working group; for example, in the third workshop representatives of all major actors/organizations in the network were present (altogether 12 participants), while in the final workshop there were only two participants present from the municipality (the coordinator of elderly care services and the coordinator of the development project). However, the core working group was active at the beginning of the project when the most important decisions about the framework and measures were taken (workshops 1-3). Another reason for variation in the working group composition was overlapping development projects; there were at least two projects running concurrently; one for productivity and another for the effectiveness of all elderly care services.

The Resulting Measurement System

Table 1 presents the results of the measurement development work in the case of Geripol. Effectiveness was divided into three perspectives: community, network and client, which facilitated the identification of measurement objects and the actual design of measures. At the community level, the aim of Geripol is to decrease the use of institutional care (and increase home care). The number of home care clients transferring to institutional care in relation to the number of residents in the municipality aged over 75 years was considered a suitable measure of effectiveness at the community level. Community perspective refers to the whole municipality, not just elderly care services. At the network level, access to services, level of cooperation in service provision, upgrading employees' competence and number of home calls conducted by the geriatrician were chosen as measurement objects. At the level of clients, measurement targets included stakeholder satisfaction (clients and family members), changes in clients' well-being, and use of emergency and unplanned institutional services.

Many of the measures, especially those related to clients' clinical aspects, were in use in municipality's elderly care. However, most of the measures required modification of current measurement and a renewal of procedures for data collection and analysis. Existing measures relate to organizations or service units in the network while the new measures of Geripol relate to the entire network. This necessitated changes to current measurement processes. The measures related to inter-organizational cooperation are new and new data collection procedures are needed in order to implement them.

Table 1. Measures of effectiveness in Geripol

Perspective	Measurement Object	Measure
Community	- Relational number of clients moved from home care to institutional care	- Number of clients moved from home care to institutional care in relation to numbers of citizens of municipality aged over 75 years
Management of the network	- Access to the service - Cooperation - Competence - Home calls	- Clients who got care in 1-7 days/all clients - Number of house calls conducted by multidiscipline team (geriatrician and nurse from home care unit) - Competence survey - Home calls/all visits to Geripol (%)
Clients	- Satisfaction of stakeholders (clients, family member etc.) - Use of hospital emergency and hospital care - Changes in client's quality of life - Changes in client's ability to function	- Stakeholder survey - Number and cost of visits to emergency room, unscheduled care days/treatment periods - Measurement of quality of life before – after treatment by Geripol - Measurement of functionality (clinical measures e.g. DRS, ALD, MAPLE) of clients before – after treatment by Geripol

At the time of the last workshop (May 2013), the purchaser had decided to implement new measures in the 2014 purchasing contracts regarding Geripol services. However, for many measures issues related to user policy were still in progress.

ANALYSIS OF FINDINGS

The case study analyzed in this paper presents a development project for designing system-level performance measures for a welfare service. The aim of the development work was to design measures related to the effectiveness of a newly launched service, the geriatric outpatient clinic Geripol. The starting point was to design measures of effectiveness for the purchaser's use, which evidenced the purchaser's powerful influence on and role in the target setting of the development project. Initially many different types of objectives related to the operations of Geripol were identified and discussed among the actors in the network. In order to clarify the big picture, and prioritize different targets, the objectives were structured into three levels of analysis based on stakeholders. The focal result of the development project relates to identifying different perspectives on effectiveness; community, client, and network management. Identifying different levels and targets related to each of them was a crucial step in deciding what to measure. This also served to clarify causal relationships between different perspectives; i.e. what should happen in service production (network management level) in order to achieve certain effects on clients, which eventually manifest as welfare effects at community level.

Effectiveness on client level was considered paramount in the measurement framework. This view was strongly supported by the purchasing manager of elderly care services, who stated that "Client is the one who steers the whole thing". Nearly all measures developed relate to the qualitative aspects of client outcomes, either directly (satisfaction, quality of life or functioning ability) or indirectly (access to services, number of home calls by multidisciplinary team, relative number of home calls). At the network level, measures were intended to support the management and development of cooperation. However, the management of the service integration process was not of the main interest in this development work, and therefore the number of measures related

to management and coordination activities was limited. On the community level, the long-term effectiveness of the Geripol services was linked to the number of home care clients entering institutional care compared to the entire elderly population of the municipality. The measures developed seemed to cover the main aspects of effectiveness at different levels of the welfare service system (cf. Crook et al., 2005; Provan & Milward, 2001), and thus providing a holistic approach for performance measurement in this context. Measures of costs and productivity were not included in these measures since they were already included in the existing measurement framework of the municipality.

Designing performance measures in the public sector is challenging (e.g. DeGroff et al., 2010; Jääskeläinen et al., 2009; Linna et al., 2010; Rantanen et al., 2007; Saunila et al., 2012). The literature suggests that one of the major challenges in designing measures at system level in public welfare services is to reach agreement on common targets and purposes of the measurement (e.g. DeGroff et al., 2010; Rantanen et al., 2007). However, consensus on the main objectives and purposes of the measurement was attained early and relatively easily in this development project, since the main target of the network's operations was clear to all participants from the beginning of the development work. Different options of measurement objects and actual measures were discussed actively in workshops and there were no major disagreements between participants about measurement. In case of minor disagreements, facilitators mediated by asking more details about different opinions, and discussions on different aspects were facilitated to reach mutual understanding of measurement. In Geripol, challenges in the design of measures related to imbalance in existing measures; many established measures were in use for some aspects, especially clinical outcomes of care, while for other aspects there were no existing measures. Those measures already in use had to be modified to the network level. For example, measures related to clients' well-being have been established in the organizations, but the level of analysis had to be modified so as to include the clients of Geripol, not all clients in the elderly care sector or in just one organization. In addition to the existing measures, the measurement system includes many new measures, especially related to network management. The implementation of those will require time, resources, and coordination.

The development work was facilitated by two researchers, and the workshops were coordinated by the research and evaluation manager of the municipality. The roles of outside facilitators and the in-house coordinator were crucial to the progress of the development work; the facilitators planned the overall content and schedules for the workshops and the coordinator scheduled them. The working team included representatives from the various organizations and professions of the network, which should ensure successful implementation of the measurement system (e.g. Jääskeläinen et al., 2013; Axelsson & Axelsson, 2006; Rantanen et al., 2007; Pekkola, 2013). However, the service network was relatively new, and roles and responsibilities were still to some extent unclear, which manifested as lack of strong coordination in the development project. Participants were also involved in other concurrent development projects (one of which related to the same theme) and this hampered the development project. Concurrent projects caused confusion among participants; e.g. some became frustrated or were too busy, which possibly undermined motivation and participation. At the end of the project, issues related to the user policy, like source data for measures and responsibilities regarding the implementation and coordination of measurement were partly still in progress.

In summary, the main findings of the measurement development project related to the identification of different levels of effectiveness and including all levels in the measurement solution. This promoted the more comprehensive measurement of the effectiveness of the welfare service system. The conceptual framework presented in the paper (Figure 1) was beneficial in the identification of the focal elements of effectiveness in this context. The division of effectiveness

into three perspectives was beneficial for performance measurement, but also for the overall performance management of the network in the future. The working group participants had common objectives for operations at the beginning of the development project, which advanced the definition of measurement objectives and measures. Workshops led by outside facilitators appear as a beneficial working method in designing performance measurement in co-operative settings such as the one examined in this study. The most focal benefits relate to the development of measurement system in cooperation among participants of the network, organized content and schedules, consultation provided by facilitators, and the facilitators' mediating role. The working group consisted of representatives from different organizations and professions, which is a prerequisite for the successful implementation and use of measures.

Clients as stakeholders of Geripol did not participate the development project, which is one of limitations of this research. However, client perspective was underlined by all participants, and the measurement solution includes various measures related to clients, e.g. direct feedback from clients, which indicates that clients' perspective is well acknowledged in measurement. Challenges related mostly to network coordination and competing development projects that participants were involved in. Presumably these factors affected the schedule of the project; the design phase alone took about 19 months, which is relatively long compared with earlier experiences in the business sector (cf. Bourne et al., 2000).

DISCUSSION AND CONCLUSION

In welfare services a more holistic approach in performance measurement and management is needed in order to avoid sub-optimization and achieve a higher level of effectiveness. While performance measurement appears promising approach in managing the effectiveness of welfare service system, research literature lacks the knowledge on how to apply it. This paper aims to contribute to the existing performance measurement literature by examining how the design of system level performance measurement targeting the measurement of effectiveness in welfare service system can be conducted. This research focuses on two focal questions in the design phase of the measurement: what to measure and how to carry out the development process in the context of inter-organizational co-operative setting.

In order to specify what to measure, this research presents a conceptual framework based on the existing literature including three perspectives on effectiveness – community, client and network management level. Whereas the earlier literature (e.g. Provan & Milward, 2001) proposes frameworks for evaluating network effectiveness at different levels, this research takes a step forward and applies such a framework in practice in the context of integrated service delivery. The framework presented in the research facilitated the identification of perspectives of effectiveness and targets related to each, which was crucial in order to prioritize and design measures. This was also helpful in illustrating the links between different perspectives; i.e., how service production (network level) affects clients, which in turn has effects at community level. Client perspective is focused in the measurement solution, either directly or indirectly, which is obvious in the context of welfare services.

To answer the question how to carry out the development process, this research implements the performance measurement design process in the context of welfare services and analyzes findings based on the earlier research literature. Earlier literature (e.g. Cunha et al., 2008; Kaplan et al., 2010) suggests that consensus on the network's target and the use of measures should be reached at the beginning of the development work. According to the results of this research, that is not a specific challenge, if the purpose and use of the measurement system are decided

beforehand, which is beneficial for the development project. Successful implementation of the measurement system requires representation and fluent cooperation of different professionals in the design phase (e.g. Axelsson & Axelsson, 2006). In this research, different organizations of the network were represented fairly well, which should enable successful implementation of the measurement system. Also, this research confirms earlier research (e.g. Axelsson & Axelsson, 2006; Bourne, 2002; Jääskeläinen, 2010; Lönnqvist, 2004; Pekkola, 2013) that contribution of outside facilitators is important in the development process. Factors identified in the earlier literature, like unestablished roles and responsibilities in the network (e.g. Hansson et al., 2010) and overlapping projects (Rantanen et al., 2007) were the main challenges that prolonged the measurement design process. Thus, the results of this research support the previous literature of the design process in many ways. However, results of this research further underlines the role of coordination; while the design process of this research entails many success factors, like the consensus of targets, outside facilitation and multidisciplinary development team, strong coordination that clarifies roles and responsibilities would have been needed in order to accomplish the task in a reasonable time.

Traditionally, the evaluation of effectiveness in welfare services is conducted by utilizing different types of economic evaluations (e.g. Nielsen & Ejler, 2008). This research has extended the scope of performance measurement approach to the measurement of effectiveness by operationalizing the concept of effectiveness in the context of integrated service delivery in welfare services and by suggesting and empirically applying a method (design process) for carrying out the measurement development. The key findings of this research relate to the identification of perspectives of effectiveness in the welfare service system and to the designing a measurement system that includes relevant perspectives. According to the present research, the more holistic measurement and management of effectiveness in welfare service system may be attained by understanding the diverse perspectives of effectiveness and taking different perspectives into account in measurement and management. The research adds to existing performance measurement literature by illustrating how the design of system level measurement that supports the integrated service delivery can be carried out in practice in the context of welfare services.

This research suggests that performance measurement approach is applicable in designing measures for effectiveness in the context of welfare service system. The conceptual framework of effectiveness and the sequential design process (workshops) are beneficial tools in designing measures of effectiveness for integrated service delivery in welfare services. Since objectives of this research were twofold: to solve a practical problem of how to measure effectiveness of a service provided in cooperation (Geripol) and to contribute to prior knowledge of performance measurement, action research was chosen as a suitable research method for the study. Action research entails many benefits, like access to real, complex decision making situations, but it has certain limitations. Research results acquired through action research are situational and contextually embedded (e.g. Coughlan & Coughlan, 2002). Thus, the results of this research, like measures developed cannot be adopted directly in any situation in measuring effectiveness in health care services or welfare services in general. Instead, both the conceptual framework and the measures need to be modified according to the strategic objectives of the network and the needs of the particular decision-making situation. The present research includes only the design phase of performance measurement, and is thus unable to add knowledge about actual implementation or use of system level performance measurement, for example, how data is collected for measures, how organizational objectives are linked to objectives of the network and how network-level measures are integrated into the management systems of individual organizations. Further research is therefore needed about the implementation and actual use of measures to support the management of effectiveness in the welfare service systems.

REFERENCES

Ahgren, B., & Axelsson, R. (2005). Evaluating integrated health care: A model for measurement. *International Journal of Integrated Care*, 5, e01–e09. PMID:16773158

Axelsson, R., & Axelsson, S. B. (2006). Integration and collaboration in public health – a conceptual framework. *The International Journal of Health Planning and Management*, 21(1), 75–88. doi:10.1002/hpm.826 PMID:16604850

Barretta, A., & Busco, C. (2011). Technologies of government in public sector's networks: In search of cooperation through management control innovations. *Management Accounting Research*, 22(4), 211–219. doi:10.1016/j.mar.2011.10.002

Bititci, U., Garengo, P., Dörfler, V., & Nudurupati, S. (2012). Performance measurement: Challenges for tomorrow. *International Journal of Management Reviews*, 14(3), 305–327. doi:10.1111/j.1468-2370.2011.00318.x

Bititci, U. S., Carrie, A. S., & McDevitt, L. (1997). Integrated performance measurement systems: A development guide. *International Journal of Operations & Production Management*, 17(5), 522–534. doi:10.1108/01443579710167230

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *Journal of Operations & Production Management*, 20(7), 754–771. doi:10.1108/01443570010330739

Bourne, M., Neely, A., Platts, K., & Mills, J. (2002). The success and failure of performance measurement initiatives, Perceptions of participating managers. *International Journal of Operations & Production Management*, 22(11), 1288–1310. doi:10.1108/01443570210450329

Brazier, J., Ratcliffe, J., Salomon, J. A., & Tsuchiya, A. (2007). *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford, New York: Oxford University Press, Inc.

Busi, M., & Bititci, U. (2006). Collaborative performance management: Present gaps and future research. *Journal of Productivity and Performance Management*, 55(1), 7–25. doi:10.1108/17410400610635471

Callender, G. (2011). Alignment of inter-agency supply chains to enhance public sector performance management. *International Journal of Productivity and Performance Management*, 60(1), 9–23. doi:10.1108/17410401111094286

Conaty, F. J. (2012). Performance management challenges in hybrid NPO/public sector settings: An Irish case. *International Journal of Productivity and Performance Management*, 61(3), 290–309. doi:10.1108/17410401211205650

Coughlan, P., & Coughlan, D. (2002). Action research for operations management. *International Journal of Operations & Production Management*, 22(2), 220–240. doi:10.1108/01443570210417515

Crook, W. P., Mullis, R. L., Cornille, T. A., & Mullis, A. K. (2005). Outcome measurement in homeless systems of care. *Evaluation and Program Planning*, 28(4), 379–390. doi:10.1016/j.evalprogplan.2005.07.011

Cunha, P. F., Ferreira, P. S., & Macedo, P. (2008). Performance evaluation within cooperate networked production enterprises. *International Journal of Computer Integrated Manufacturing*, 21(2), 174–179. doi:10.1080/09511920701607907

DeGroff, A., Schooley, M., Chapel, T., & Poister, T. H. (2010). Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning*, 33(4), 365–372. doi:10.1016/j.evalprogplan.2010.02.003 PMID:20303176

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. New York: Oxford University Press Inc.

Evans, J. M., & Ross Baker, G. (2012). Shared mental models of integrated care: Aligning multiple stakeholder perspectives. *Journal of Health Organization and Management*, 26(6), 713–736. doi:10.1108/14777261211276989 PMID:23252323

- Flatau, P., & Zaretsky, K. (2008). The Economic Evaluation of Homelessness Programmes. *European Journal of Homelessness*, 2(12), 305–320.
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., & Neely, A. et al. (2007). Towards a definition of a business performance measurement system. *International Journal of Operations & Production Management*, 27(8), 784–801. doi:10.1108/01443570710763778
- Gummesson, E. (2000). *Qualitative Methods in Management Research*. Thousand Oaks, California: Sage Publications Inc.
- Halligan, J., Sarrico, C. S., & Lee Rhodes, M. (2012). On the road to performance governance in the public domain? *International Journal of Productivity and Performance Management*, 61(3), 224–234. doi:10.1108/17410401211205623
- Hansson, J., Qvretveit, J., Askerstam, M., Gustafsson, C., & Brommels, M. (2010). Coordination in networks for improved mental health service. *International Journal of Integrated Care*, 10(25), 1–9. PMID:20922065
- Jääskeläinen, A. (2010). *Productivity Measurement and Management in large Public Service Organizations*. TUT Publication 927. Tampere: Tampere University of Technology.
- Jääskeläinen, A., & Lönnqvist, A. (2009). Designing operative productivity measures in public services. *VINE: The journal of information and knowledge management systems*, 39(1), 55–67.
- Jääskeläinen, A., & Sillanpää, V. (2013). Overcoming challenges in the implementation of performance measurement: Case studies in public welfare services. *International Journal of Public Sector Management*, 26(6), 440–454. doi:10.1108/IJPSM-12-2011-0014
- Jansen, E. P. (2008). New Public Management; perspectives on performance and the use of performance information. *Financial Accountability & Management*, 24(2), 169–191. doi:10.1111/j.1468-0408.2008.00447.x
- Jones, K., Colson, P., Valencia, J. D., & Susser, E. (1994). A Preliminary Cost Effectiveness Analysis of an Intervention to Reduce Homelessness Among the Mentally Ill. *The Psychiatric Quarterly*, 65(4), 243–256. doi:10.1007/BF02354302 PMID:7831412
- Kaplan, R. S., Norton, D. P., & Rugelsjoen, B. (2010). Managing Alliances with the Balanced Scorecard. *Harvard Business Review*, (January-February): 114–120.
- Kulmala, H., & Lönnqvist, A. (2006). Performance measurement of networks: Towards a non-financial approach. *International Journal of Networking and Virtual Organizations*, 3(3), 299–316. doi:10.1504/IJNVO.2006.010953
- Kurunmäki, L., & Miller, P. (2011). Regulatory hybrids: Partnerships, budgeting and modernising government. *Management Accounting Research*, 22(4), 220–241. doi:10.1016/j.mar.2010.08.004
- Laihonen, H., Jääskeläinen, A., & Pekkola, S. (2014). Measuring performance of a service system – from organizations to customer-perceived performance. *Measuring Business Excellence*, 18(3), 73–86. doi:10.1108/MBE-08-2013-0045
- Linna, P., Pekkola, S., Ukko, J., & Melkas, H. (2010). Defining and measuring productivity in the public sector: Managerial perceptions. *International Journal of Public Sector Management*, 23(5), 479–499. doi:10.1108/09513551011058493
- Lönnqvist, A. (2004). *Measurement of Intangible Success Factors: Case Studies on the Design, Implementation and Use of Measures*, TUT Publication 475. Tampere: Tampere University of Technology.
- Lönnqvist, A., & Laihonen, H. (2012). Welfare service system productivity: The concept and its application. *International Journal of Productivity and Performance Management*, 61(2), 128–141. doi:10.1108/17410401211194644
- Martin, L. L., & Kettner, P. M. (2010). *Measuring the Performance of Human Service Programs*. Thousand Oaks, California: Sage Publications, Inc.

- Murray, C. J. L., & Frenk, J. (2000). A framework for assessing the performance of health system. *Bulletin of the World Health Organization*, 78(6), 717–731. PMID:10916909
- Neely, A., Gregory, M., & Platts, K. (2005). Performance measurement system design. A literature review and research agenda. *International Journal of Operations & Production Management*, 25(12), 1228–1263. doi:10.1108/01443570510633639
- Nielsen, S. B., & Ejler, N. (2008). Improving Performance? Exploring the Complementarities between Evaluation and Performance Management. *Evaluation*, 14(2), 171–192. doi:10.1177/1356389007087538
- Osborne, S. P. (2006). The New Public Governance? *Public Management Review*, 8(3), 377–387. doi:10.1080/14719030600853022
- Packard, T. (2010). Staff Perceptions of Variables Affecting Performance in Human Service Organizations. *Nonprofit and Voluntary Sector Quarterly*, 39(6), 971–990. doi:10.1177/0899764009342896
- Parung, J., & Bititci, J. (2008). A Metric for Collaborative Networks. *Business Process Management Journal*, 14(5), 654–674. doi:10.1108/14637150810903048
- Pekkola, S. (2013). *Performance Measurement and Management in a Collaborative Network*, Acta Universitatis Lappeenrantaensis 534. Lappeenranta: Lappeenranta University of Technology.
- Porter, M. E. (2010). What is Value in Health Care? *The New England Journal of Medicine*, 363(26), 2477–2481. doi:10.1056/NEJMp1011024 PMID:21142528
- Provan, K. G., & Milward, H. B. (2001). Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational Networks. *Public Administration Review*, 61(4), 414–423. doi:10.1111/0033-3352.00045
- Qvretveit, J. (2002). How to run an effective improvement collaborative. *International Journal of Health Care Quality Assurance*, 15(5), 192–196. doi:10.1108/09526860210437403
- Rantanen, H., Kulmala, H. I., Lönnqvist, A., & Kujansivu, P. (2007). Performance measurement systems in the Finnish public sector. *International Journal of Public Sector Management*, 20(5), 415–433. doi:10.1108/09513550710772521
- Saunila, M., Ukko, J., Rusila, P., & Rantanen, H. (2012). Challenges of performance measurement and management in public healthcare services. *World Review of Entrepreneurship, Management and Sust. Development*, 8(2), 181–195.
- Sillanpää, V. (2011). Performance measurement in welfare services: A survey of Finnish organisations. *Measuring Business Excellence*, 15(4), 62–70. doi:10.1108/13683041111184116
- Strandberg-Larsen, M., & Krasnik, A. (2009). Measurement of integrated healthcare delivery: A systematic review of methods and future research directions. *International Journal of Integrated Care*, 9(2), e01–e10. PMID:19340325
- Westbrook, R. (1995). Action research: A new paradigm for research in production and operations management. *International Journal of Operations & Production Management*, 15(12), 6–20. doi:10.1108/01443579510104466
- Wistow, G., & Dickinson, H. (2012). Integration: Work still in progress. *Journal of Health Organization and Management*, 26(6), 676–684. doi:10.1108/14777261211275881 PMID:23252320
- Yin, R. K. (2009). *Case Study Research, Design and Methods* (4th ed.). Thousand Oaks, California: Sage Publications, Inc.

Virpi Sillanpää received her MSc in Economics and works as a researcher at the Performance Management team at Tampere University of Technology. She is currently engaged in doctoral studies related to the measurement of welfare services.