

Research and Theory

Consensus group sessions: a useful method to reconcile stakeholders' perspectives about network performance evaluation

Marie-Eve Lamontagne, École de réadaptation [School of Rehabilitation], University of Montréal, P.O. Box 6128, Station Centre-ville, Montréal, QC H3C 3J7, Canada

Bonnie R. Swaine, École de réadaptation [School of Rehabilitation], University of Montréal, P.O. Box 6128, Station Centre-ville, Montréal, QC H3C 3J7, Canada

André Lavoie, Unité de traumatologie—urgence—soins intensifs, Centre de recherche FRSQ du CHA (Centre Hospitalier affilié universitaire de Québec) [Unit Traumatology—Emergency Medicine—Intensive Care, CHA, FRSQ Research Centre], 1401, 18th Street, Québec, QC G1J 1Z4, Canada

François Champagne, Department of Health Administration, University of Montréal, P.O. Box 6128, Station Centre-ville, Montréal, QC H3C 3J7, Canada

Anne-Claire Marcotte, Agence d'Évaluation des Technologies et des Modes d'Intervention en Santé [Agency for the Evaluation of Health Technologies and Interventions], 2021, av. Union, bur. 10.083, Montréal, QC H3A 2S9, Canada

Correspondence to: Bonnie R. Swaine, School of Rehabilitation, Faculty of Medicine, University of Montréal, 7077, Avenue du Parc, C.P. 6128 Succ. Centre-ville, Montréal, Québec H3C 3J7, Canada, Phone: +1-514-343-7361, Fax: +1-514-343-2105, E-mail: bonnie.swaine@umontreal.ca

Abstract

Background: Having a common vision among network stakeholders is an important ingredient to developing a performance evaluation process. Consensus methods may be a viable means to reconcile the perceptions of different stakeholders about the dimensions to include in a performance evaluation framework.

Objectives: To determine whether individual organizations within traumatic brain injury (TBI) networks differ in perceptions about the importance of performance dimensions for the evaluation of TBI networks and to explore the extent to which group consensus sessions could reconcile these perceptions.

Methods: We used TRIAGE, a consensus technique that combines an individual and a group data collection phase to explore the perceptions of network stakeholders and to reach a consensus within structured group discussions.

Results: One hundred and thirty-nine professionals from 43 organizations within eight TBI networks participated in the individual data collection; 62 professionals from these same organisations contributed to the group data collection. The extent of consensus based on questionnaire results (e.g. individual data collection) was low, however, 100% agreement was obtained for each network during the consensus group sessions. The median importance scores and mean ranks attributed to the dimensions by individuals compared to groups did not differ greatly. Group discussions were found useful in understanding the reasons motivating the scoring, for resolving differences among participants, and for harmonizing their values.

Conclusion: Group discussions, as part of a consensus technique, appear to be a useful process to reconcile diverging perceptions of network performance among stakeholders.

Keywords

integrated health care delivery, network, evaluation, performance, brain injury

Introduction

Over the last decades, health care systems have faced important challenges in dealing with pressures to increase the performance of the services they offer while the allocated resources to offer such care have remained stable or have decreased. Integration has been presented as an efficient solution to improve service delivery by tackling fragmentation and efficiency issues. One of the many forms of service integration is inter-organizational networks, which could be defined from a sociometric perspective as a set of three or more organizations linked together using different strategies [1, 2]. Other theorists, however, suggest that it is the mode of governance, neither hierarchical or competitive, that defines a network [2, 3]. Despite their apparent simplicity, networks are complex organisations that vary greatly on many aspects. Indeed, networks can involve various organizational levels of the organization, such as administration, governance, caseload, management, etc. [4, 5]. They may be created at various levels, bringing governmental agencies, regional or local facilities or programs [4] to work together using one or many integration strategies, such as communication, resource sharing, joint programs, client transfer agreements and so on [1]. Depending on the intensity of the links created, networks can be cooperative, coordinated or collaborative [6, 7], the latter being the most intensive and the most demanding. According to the theoretical perspective fostered, networks can also be thought of as a source of knowledge and information, a system of power regulation or a social structure created by environmental tensions [2]. A very common perspective of networks in the field of health services is that of the services perspective, where network partners work together to create a seamless continuum of care. Because they are viewed by many service professionals, policy makers and researchers as a potential integration strategy that could reduce service fragmentation and improve the coordination of services and effectiveness [8], networks have become a popular way to reconfigure health service delivery systems [9–11] found in many countries. In addition, the challenging context in which health care systems must operate has generated a growing commitment to the evaluation of health services. Indeed, programs,

facilities and systems are increasingly required to demonstrate their performance. The rising popularity of networks and the renewed importance of health services evaluation have bred new interest in the evaluation of integrated networks.

The evaluation of integrated networks is complicated by the fact that the concept of network performance is poorly defined [3]. This is due in part to the fact that the advantages of networks have been predominantly studied with regard to their effect on individual organizations, and not with regard to the whole network [3, 10]. To date, network performance has been mainly appraised using performance indicators traditionally used for individual organizations. At best, this may lead to partial and inaccurate evaluations of network performance [10, 11]. The lack of clarity about the concept of network performance is likely to induce difficulties in network performance evaluation, and consequently limits our ability to use evaluation results to improve network performance [12].

In addition, the evaluation of integrated networks is complicated because networks involve many organizations that may have different values, cultures and mandates. Consequently, network members could have different perspectives about what constitutes the performance of their network [11, 13, 14]. Moreover, Provan et al. [15] stated that network stakeholders tend to see the network and its performance from the perspective of their own organization. Huxham [14] also mentioned that the variety of organizational and individual agendas that are present in collaborative situations make it difficult to agree in practice. This author further suggests that when partners do not completely agree on a shared purpose, they may not be able to agree on the next steps of a collaborative process [16]. Indeed, if they do not agree on what constitutes new network performance, network partners could be in conflict on the evaluation process and indicators and interpret differently the evaluation results. A lack of a common vision risks disintegrating the collaborative advantage of network.

To address various stakeholders' perspectives of organizational performance, the multiple-constituency approach to evaluation has been proposed [17]. This approach states that the multiple constituencies

or stakeholders of an organization could hold different perspectives about what constitutes its performance, and that it is unjustified or arbitrary to choose one perspective over another for evaluation purposes. According to the multiple-constituency approach, these perspectives should be incorporated into a larger vision reflecting the sum of the evaluative criteria applied by the various stakeholders involved. Several studies have used this approach to build performance frameworks for various organizations [12, 18–20], most of them having found differences in perceptions and priorities among different types of stakeholders. However, the use of this approach within a performance evaluation process presents two major weaknesses. The first weakness is that the multi-constituency approach does not explicitly explain how to deal with divergent perceptions. Without further interaction among stakeholders, each group maintains its own different understanding of performance despite being embedded within a larger framework. Deciding upon what to include in a performance evaluation framework, based solely on a multi-constituency approach, will thus likely fail to harmonize the standards and values of the different stakeholder groups and may be unsuccessful in promoting collaboration and mutual trust between network members [21].

A second challenge associated with using the multi-constituency approach for performance evaluation is that the methods frequently used within this approach (written questionnaires) provide only a limited and thus superficial understanding of the underlying rationale for the differing opinions of stakeholders. Moreover, responding to a questionnaire does not provide stakeholders with an opportunity to share and to discuss their preferences to better understand their partners' vision. Many researchers [20, 22, 23] have emphasized the need to go beyond simply reporting differences in stakeholder perceptions and to engage stakeholders in meaningful discussions to arrive at a common understanding of an organization's performance. This may be particularly important when integrated network links partners from different settings with different values and priorities [24]. However, few authors have proposed ways to do so. Indeed, Zinn and colleagues [25] repeated the steps of a Delphi technique to obtain consensus within homogenous stakeholder groups regarding laboratory management performance indicators. They found the technique useful for reaching a consensus within groups of stakeholders of a similar type (e.g. hospital executives, managed care executives, referring physicians, laboratory regulators and laboratory managers). However, they did not explore whether the technique was effective in reconciling the perceptions across

various stakeholder types. The Delphi technique has certain benefits (e.g. anonymity and consultation of persons over large distances) but it does not provide opportunities for meaningful discussions between stakeholders [26]. Discussion could be useful especially when the subject under study is complex and abstract [26], such as organizational performance. Such subjects might thus be better explored through methodology involving group discussions. Studies of collaboration highlight the importance of a drafting process that is highly participatory and involves key stakeholders [27].

The creation of a common vision between partners sets the foundations for collaborative efforts and enhances the collaborative advantage of network. However, traditional approaches integrating multi-constituency perspectives fail to reconcile partners' visions and involve them only superficially in the process. To gain a better understanding of what constitutes network performance and to improve the performance evaluation process, there is a need to explore innovative approaches to reconcile stakeholders' perceptions while giving them an opportunity to discuss and share ideas. Falling short of this may result in fragmented network efforts, disinterest of network partners with regard to the vision, evaluation process and evaluation outcome, they deem as not responsive to their reality.

TRIAGE, a consensus technique, could be a viable method to create a shared vision in a participatory way. The general goal of this study was to explore whether TRIAGE group discussions could reconcile the perspectives of network members with regard to the importance of performance dimensions. Specifically, we 1) determined the initial degree of consensus existing among individual TBI network organization members with regard to importance of performance dimensions for performance evaluation, 2) explored the usefulness of consensus group sessions to reconcile these perceptions, by comparing the importance paid to the different dimensions of performance using questionnaires and group discussion.

Method

We conducted the present study from November 2006 to November 2008 during the accreditation process of network organizations providing services to persons with traumatic brain injury in the province of Québec (Canada). We used the TRIAGE consensus technique [28, 29], a method that combines an individual data collection phase using questionnaires with a group data collection phase involving group consensus sessions.

Procedures

In a previous study [12], we used a questionnaire to survey the representatives of 46 individual organizations (i.e. acute care facilities, rehabilitation facilities, regional health authorities, the accreditation body and the provincial health ministry). We asked participants to determine the relative importance of the 16 dimensions included in the EGIPSS (*Evaluation Globale et Intégrée de la Performance des Systèmes de Santé*) performance framework (Figure 1) [23] for the evaluation of the performance of TBI network using a scale where 0% indicates not important at all and 100% indicates extremely important. For the present study, we re-analyzed these survey data (excluding data from the accreditation body and the ministry) by compiling the results for each specific network, where a network was composed of an acute care facility, one or more rehabilitation facilities and regional health authorities. These data were considered the result of the individual data collection phase of TRIAGE.

Then, for the group data collection, the accreditation process provided an opportunity to bring together network members who are otherwise geographi-

cally spread across each of the regions. The group participants were the clinical coordinators and the managers from provincially-funded TBI programs. A group session was organized for each network. The participants completed socio-demographic data and consents forms at the beginning of the session. The group animator then reminded participants of the study goals and proposed an agenda for the session. The discussions began with the dimensions of the EGIPSS framework for which differences in the importance scores attributed in the individual data collection phase by members within the same network were inferior to 20%. For these consensual dimensions, a consensual score, based on the average of the importance scores attributed by the individual organizations, was attributed.

In contrast, dimensions having a score difference $\geq 20\%$ were discussed in depth. As per the TRIAGE technique, the names of these dimensions were written on cardboard sheets and stuck on the wall under the Grouping section for viewing by the group. Participants were invited to explain why their individual organization teams had rated the importance of a dimension the way they did. Then, as network members, they discussed and agreed upon the importance score *the network*



Source: Adapted from Champagne F, Contandriopoulos AP, Picot-Touché J, Béland F, Nguyen H : Un cadre d'évaluation de la performance des systèmes de services de santé : Le modèle EGIPSS. Conseil de la santé et du Bien-être, Gouv. du Québec. www.csbe.gouv.qc.ca

Figure 1. EGIPSS performance framework.

should attribute to the dimension. When consensus was attained, the dimension name (and new importance score if different from zero) was placed on the wall under the Selection heading. Dimensions rejected as being unimportant to the network (i.e. scored as having zero importance) were placed in a Garbage section. Dimensions for which consensus could not be reached quickly were put in a Fridge section for further discussion at the end of the session. If a consensus could not be reached, even after these additional discussions, the dimensions could be placed in a Veto section and subsequently submitted to an external expert committee. Using this technique, a final list of dimensions, each with a consensus-based importance score, was created.

Participants

Eight TBI networks existed in Québec at the time of the study. These networks linked trauma centers designated to provide specialized trauma care with facilities providing in- and/or out-patient rehabilitation services and with regional health authorities. Five of these the networks included each three organizations and operated on a regional basis (e.g. member organizations were all located in the same socio-demographic region). The three other networks operated on a larger inter-regional level ensuring specialized trauma care and/or in-patient rehabilitation services availability to regions of Québec not having these services in their region. One inter-regional network was comprised of six organizations, while two others linked 11 and 13 different organizations. Members of the accreditation body and representatives from the health ministry were excluded from the group sessions since they were technically members of each network. The research Ethics Committees of each organization approved the study prior to data collection.

Analysis

The minimum and maximum importance scores attributed to the dimensions of the EGIPSS framework by the individual members of each network were compiled, and the differences between these scores were calculated to determine the dimensions to be discussed during the group sessions. For each dimension, we computed the median importance score and its mean rank using individual organisation and group data. Mann–Whitney tests were used to examine differences between individual and group data for the 16 dimensions. Pearson χ^2 and Mann–Whitney tests were also used to compare the characteristics of the participants involved in the individual and group data collection phases. Statistical analyses were conducted

with SPSS 16.0 software, using a p-value of 0.05 as the significance level. The consensus group sessions were taped and transcribed, and the verbatim was analysed using Nvivo 8.0 software.

Results

Participants' characteristics

One hundred and thirty-nine professionals from acute care facilities (25%), rehabilitation facilities (65%) and regional health authorities (10%) participated in the individual data collection (questionnaires). Participants had a mean professional experience of 17 years (SD 9.7) and reported working in their current position for an average of eight years (SD 6.1). All existing networks (n=8) agreed to participate in the study and 62 persons from these eight networks participated in the group sessions: 27% were from acute care facilities, 52% from rehabilitation facilities and 21% represented regional health authorities. Group participants had a mean professional experience of 18 years (SD 9.7) and reported working in their current position for eight years (SD 8.8). The participants of the individual and group data collection phases were similar in terms of work experience and types of organization.

Consensus observed through questionnaire use

Table 1 presents the minimum and maximum importance scores attributed by the individual participants within a network to the dimensions of the EGIPSS framework; differences $\geq 20\%$ are highlighted in gray. Overall, the differences in the importance scores within the networks ranged from 0% to 100%, with a mean range of 33.7%. In fact, for the majority of dimensions, the observed individual importance score differences exceeded 20% (66%, corresponding to 85/[8 networks \times 16 dimensions]).

With regard to the variability of the importance scores across networks, we found the networks had on average 10.2/16 dimensions for which the score ranges were $\geq 20\%$. The extent of consensus on importance scores varied according to the network examined. For the inter-regional network A, there was a total absence of agreement in scores, while for regional networks G and H, the majority of the performance dimensions scored within the 20% range. Small networks were more consensual than larger ones.

The extent of consensus about importance scores also varied according to the performance domain.

Table 1. Minimum and maximum importance scores attributed by the individual organization members of each network

		Networks (n=8)								
Domains of the EGISS framework (n=4)	Type of network Number of organization members	A	B	C	D	E	F	G	H	No. of networks with range of scores ≥20
		Inter-regional 14	Inter-regional 13	Inter-regional 11	Regional 3	Regional 3	Regional 3	Regional 3	Regional 3	
Adaptation	Ability to adapt to requirements and tendencies	20-90	40-100	20-100	75-100	60-80	60-90	40-55	50-95	7
	Ability to innovate and transform	0-100	80-100	80-100	80-90	50-85	60-80	60-80	60-90	7
	Ability to mobilize community support	30-100	30-95	40-100	60-100	80-90	80-90	50-85	20-85	6
	Ability to adapt and meet the client's needs	60-100	80-100	80-100	80-100	90-100	80-95	80-95	60-85	5
	Capacity to acquire resources	55-100	20-95	35-100	85-100	80-100	75-90	90-90	0-80	5
Value maintenance	Capacity to attract the clientele	70-100	20-100	95-100	80-100	100-100	95-100	85-90	85-95	3
	Consensus with fundamental values	15-100	50-90	35-100	75-100	85-100	80-95	50-85	75-100	6
Production	Collaborative climate	15-100	80-95	35-100	65-90	80-100	60-95	75-90	75-80	5
	Productivity	0-100	20-80	30-100	20-95	50-80	60-85	20-70	85-90	7
	Quantity of care and services	10-100	20-60	10-100	20-90	20-75	50-80	40-70	90-95	7
	Quality	70-100	80-100	0-100	95-100	40-100	90-95	85-90	90-100	4
Goal attainment	Continuity	70-100	50-100	0-100	95-100	90-100	90-95	70-90	90-100	4
	Satisfaction of clients and partners	0-100	60-85	40-100	80-90	85-90	50-95	55-80	80-90	5
	Effectiveness	60-100	80-100	50-100	50-90	90-100	65-95	80-80	90-90	5
	Efficiency	70-100	40-100	90-100	80-100	80-100	90-95	85-95	70-95	5
Equity	Equity	30-100	70-95	45-100	80-95	90-100	70-95	75-80	90-100	4
	Number of dimensions with a range of scores ≥20	16	15	14	10	8	8	7	7	85/128

Three domains of the EGIPSS model, i.e. the Adaptation, Values maintenance and Production, had similar proportions (31.3%) of dimensions with scores with ranges of <20% (results not included in Table 1). In contrast, the Goal attainment domain showed greater consensus with 40.6% of its corresponding dimensions with ranges of 20% or less. The extent of consensus also varied according to the performance dimension under consideration. Indeed, four dimensions were attributed a wide range of scores by 7/8 networks: Ability to adapt to requirements and tendencies, Ability to innovate and transform, Productivity, and Quantity of care and services. In contrast, three dimensions (Efficiency, Quality and Continuity) were attributed a range of scores $\geq 20\%$ by only half (4) of the networks and only one dimension (Capacity to attract the clientele) was similarly scored by the majority of the networks.

Consensus observed through group sessions

Eight consensus group sessions were held, each lasting on average 1.6 hours. Eighty-five discussions took place about the different dimensions. On nine occasions, because the group deemed them unimportant, dimensions were placed in the Garbage and received an importance score of 0%. Four groups rejected Quantity of care and service, Productivity was

rejected by three, and Quality and Effectiveness by one each. On eight occasions, at least one dimension was temporarily placed in the Fridge section. However, after discussion, the groups were able to establish a consensus about their importance such that no dimension was placed in the Veto section. At the end, consensus was established for 100% of the dimensions for each network.

Comparison of the consensual scores

Table 2 details, in decreasing order, the median importance scores after group discussions, the inter-quartile ranges (IQR) and the mean rank of the EGIPSS dimensions obtained in the individual and group data collection phases. The median importance scores attributed to the dimensions via the individual survey varied from 60% to 100% (median of 85%), while those attributed through consensus varied from 20.0% to 96.7% (median of 86.6%). Overall, the differences between the scores attributed in the individual and group data collection phases were not statistically significant ($Z = -0.88, p = 0.378$).

Five dimensions received higher importance scores during consensus group sessions as compared to the individual data collection phase. The Capacity to attract the clientele, Continuity, and Efficiency dimensions were consistently scored as the three most important dimensions, both during the individual

Table 2. Median importance scores of EGIPSS dimensions from individual and group production phase

Dimensions of EGIPSS framework (n=16)	Individual production phase		Group production phase	
	Median importance score (%) [IQR]	Mean ranks	Median importance score (%) [IQR]	Mean ranks
Capacity to attract the clientele	100 [90–100]	4.3	96.7 [93.4–99.7]	3.2
Continuity	95 [85–100]	5.4	96.6 [94.8–100]	2.0*
Efficiency	92.5 [80–100]	6.0	94.2 [90–95.4]	4.9
Quality	90 [80–100]	6.1	92.3 [85–95.8]	5.9
Ability to adapt and meet the client’s needs	90 [80–100]	6.2	91.7 [86.7–95.7]	5.8
Equity	90 [80–95]	7.1	90 [88–94.2]	6.8
Consensus with fundamental values	85 [75–95]	8.9	89.2 [84.2–90]	7.7
Satisfaction of clients and partners	80 [80–90]	9.5	88.2 [81.7–92.5]	7.4
Effectiveness	90 [80–90]	8.1	87 [85–95]	7.7
Capacity to acquire resources	90 [80–95]	8.2	85 [70–88.8]	9.1
Ability to innovate and transform	80 [80–90]	9.6	81.7 [77.5–85]	10.9
Collaboration climate	90 [80–95]	8.8	80.7 [79.2–86.3]	10.3
Ability to mobilize community support	80 [60–90]	10.4	72.5 [52.5–81.7]	11.9
Ability to adapt to requirements and tendencies	70 [55–80]	12.3	70 [49.2–80]	12.8
Productivity	80 [50–85]	11.4	25**[0–60]	13.9
Quantity of care and services	60 [40–75]	13.1	20***[0–70]	14.5
Total	85 [75–95]	–	86.6 [76.6–93.3]	–

* $Z = -3.61, p < 0.001$.

** $Z = -2.45, p = 0.01$.

*** $Z = -2.05, p = 0.04$.

and group production phases. The scores attributed through consensus group sessions were slightly lower for nine of the 16 dimensions. The Ability to adapt to requirements and tendencies and Quantity of care and services remained among the least important dimensions in the two data collection phases. The differences in importance scores were statistically significant only for the Productivity and the Quantity of care and services dimensions, for which the importance scores attributed by participants dropped dramatically during the group process. With regard to the mean ranks indicating the relative importance of performance dimensions, the group process did not greatly influence the order of importance of the dimensions, with the exception of the Continuity dimension, which was scored as relatively more important during the group sessions.

It is noteworthy that the dimensions with the highest importance scores (e.g. Capacity to attract the clientele, Efficiency, Quality, etc.) (Table 2) were also the most consensual (Table 1).

Explanations provided through group discussions

Group discussions provided the participants with an opportunity to justify their individual scoring and to agree as a group on a new importance score. Many of the general comments pertained to the desired balance between the integration and the differentiation of network partners. Indeed, while agreeing that some degree of integration is required to provide quality services, network members wish to remain distinct organizations with their own culture, values, and intervention methods.

When discussing the dimensions related to the Adaptation domain, participants perceived Ability to adapt and meet the client's needs as the only one fundamental to a network. The other five dimensions of the domain were seen as additional activities going beyond the fundamentals of a network. Because they require considerable resources and energy to be carried out, participants highlighted the need to carefully consider the impact of engaging in activities requiring the network to adapt and transform. Participants commented on the potential for improvement intrinsic to integrated organizations. Indeed, working together gives network participants opportunities to examine their practice through the eyes of their partners, to share and to exchange new knowledge likely to improve their network activities. Some network participants attributed less importance to the dimensions deemed as encompassing relations with the external environment, such as Capacity to acquire resources

and Ability to mobilize community support. This was because they felt they had little control over their external environment and they did not want to be labeled as less performing if, for instance, they failed to mobilize community support: "One could put a lot of energy into trying to mobilize partners in the community, but at the end it is the partners who decide whether or not to provide their support". The external environment was seen as more of a constraint than an opportunity for adaptation and transformation. However, a few participants mentioned that the difficulties encountered in the external environment are likely to bring individuals, organizations and networks to 'think outside the box' and to find creative solutions to improve their performance.

With regard to dimensions included in the Value maintenance domain, group participants explained that the Collaborative climate could facilitate or impede partnerships at the political or managerial levels, but in the end, it has a small effect on the quality of care provided to patients. They also mentioned that the Consensus with fundamental values dimension was seen as a prerequisite to any collaboration. However, because they felt it is already present in the current public health system, it was not deemed to be a sensitive indicator of network performance.

With regard to the dimensions belonging to the Production domain, Continuity was described as a very fundamental dimension of the performance of the TBI networks. Indeed, participants indicated that the collaborative links and coordination established between network participants aims primarily to enhance the coherence of services (e.g. theoretical frameworks, discharge and admission criteria, etc.) and their seamlessness (e.g. absence of delays between inter-facility transfers, absence of service interruption). The Quality of care and services dimension was seen as important, but was perceived as being embedded in other dimensions, such as Continuity and Satisfaction of clients and partners. Members of one network decided to give a score of 0% (Not important at all) to this dimension because they believe that the evaluation of quality is the exclusive responsibility of an organization, and not of a network. The two others dimensions, Productivity and Quantity of care and services, suffered a considerable drop in their mean importance scores when discussed in groups. Productivity was also described as being under the strict responsibility of individual facilities not of the network's, while Quantity of care and services was considered as useless if not associated with other measures such as the number of clients seen.

Finally, with regard to the dimensions related to the Goal attainment domain, all participants agreed that

it is fundamental for the network to reach its goals. Furthermore, participants explained that Efficiency is a more important performance dimension to appraise than Effectiveness, (“*You do what you have to do*” versus “*You do what you have to do using less resources*”). The former dimension was perceived as providing more information, while the latter was perceived as a manifestation of conformity toward authority expectations. Since they perceived Effectiveness as embedded in the Efficiency dimension, one network decided to reject (attribute a score of 0% importance) the Effectiveness dimension to avoid redundancy in the evaluation. The Satisfaction of clientele and partners was described as central to client-centered care, but participants expressed concerns about the subjectivity of patient satisfaction evaluations as they can be influenced by external elements or unrealistic expectations. The notion of Equity raised many philosophical considerations. Group participants explained that the very creation of integrated health care for persons with TBI could lead to an inequity in the service provision for other patients; TBI networks require many resources possibly depleting them for remaining patient populations.

Discussion

In this study, we first observed that when obtained through individually answered questionnaires, the extent of consensus about the importance of performance dimensions is low and network members have different perceptions about the importance of at least 50% of the performance dimensions. This suggests that any methodology that does not involve interactions between stakeholders is not likely to produce consensual results, and that a performance framework based on survey results could fail to promote a shared vision among network stakeholders. We also observed that the extent of consensus varied across networks and according to the dimensions examined. Several studies using a multi-constituency approach involved participants from multiple settings [18, 20, 25]. However, to our knowledge, our study is the first to examine the variability in the extent of consensus across integrated settings (e.g. networks). Larger networks exhibited more differences in stakeholder perceptions, while smaller ones were typically (but not systematically) more consensual. This is not surprising since networks linking a large number of participants are more likely to include heterogeneous organizations. Our findings thus support Huxham’s [14] theory, in that the complexity of partnerships seems to influence mutual understanding. However, it is also possible that participants of smaller networks had the opportunity to be in closer proximity to each other and

developed stronger links. The small size of the network could thus facilitate the adoption of a common vision. Even in smaller networks, we observed an absence of consensus for at least half of the dimensions surveyed. In combination, these results suggest that consensus group sessions could be beneficial for organizations of all sizes, but especially for larger ones, regrouping many members potentially holding different perceptions of what is important to consider for performance evaluation.

The results of the survey also indicate that the dimensions deemed more important by network participants are more consensual than others. This result suggests that the participants intuitively agree on the higher importance of a core set of dimensions often qualified as fundamental to the networks when discussed in groups. The fact that a high importance was paid to dimensions such Ability to attract the clientele (Accessibility) and Continuity supports the idea that the performance of an integrated network of care is different from the performance of a single organization [7, 10, 11], but shares the goal of many integration initiatives [4]. Because the concept of network performance is relatively new and yet not fully shared by all network members, its specificity should be made explicit to all stakeholders before any evaluation process takes place to avoid different visions tainted by the characteristics of a single organization (rather than of a whole network).

Another important result of our study is that overall, consensus group sessions did not influence the quantitative importance attributed to performance dimensions by the member participants. Given the small changes in importance scores provided by group participants, one might question the usefulness of consensus group sessions over the simple computation of median importance scores compiled using an individual data collection phase. We argue that the added value of consensus group sessions is evident in the discussions made possible between participants. Indeed, qualitative analysis of the group discussions showed that participants expose the reasons that motivated the initial and group scoring. The exchanges help other network members understand their visions and it helps sustain discussions leading to a shared understanding of network performance. Such harmonization of values and standards helps increase mutual trust and confidence among partners, which are deemed essential to improve collaboration [6, 14] and to increase network performance. A second important added value of the consensus group session is that it represents a bottom-up approach to evaluation, in that it involves stakeholders in the design of the evaluation process. Because group discussion takes into account the values and interest of networks

partners, they are likely to improve the responsiveness and the usefulness of the evaluation process as well as stakeholders' adherence to the resulting performance framework. Consensus group discussion can also help to empower network participants who usually have less power in the network [14], by giving them an opportunity to take part in the decision about the network performance. This could facilitate the 'collaborative advantage' and neutralize the 'collaborative inertia' that are often observed in collaborative efforts [14], such as networks.

In general, the results suggest that consensus group sessions are a useful step in the process of developing a TBI network performance evaluation framework and that the reconciliation of stakeholder perspectives could be a logical complement to multi-constituency approach studies. They could be used in any type of network, but would be most suitable for service networks, such as the one studied here and they may be more useful to participants who already have minimal experience of working together [14]. Consensus group sessions may be useful to clarify network performance frameworks, but could also sustain goal or process clarification. As such, the TRIAGE technique may represent a tool for integrated care, to improve a common vision and language. For an example, it could help practitioners who work collaboratively in a joint program to explore their perceptions of the program's goals, and to discuss the ones to prioritize. It constitutes a collaborative and reflexive tool that taps into benefits from the multiple perspectives and empowers the participants into establishing a shared construction of a common vision. It could be beneficial to the effectiveness and to the climate of a network, two important dimensions of network performance [12].

Group sessions allow participants to decide upon objectives they deem important, while using minimal time and human resources. With regard to the effectiveness of TRIAGE, the technique allowed the participants in each network to discuss their perceptions and to agree upon the importance they attributed to the dimensions discussed. This surpasses the 80% cut-off typically used in consensus methods, such as the Delphi technique. TRIAGE, however, requires travelling for some participants, which could be difficult when a network is geographically spread out. Group discussions could be limited by time restrictions or influenced by the level of expertise or the personality of participants.

We acknowledge that this research has some limitations. We used an arbitrary cut-off of 20% to determine the dimensions requiring discussion during the sessions. Using another cut-off may have led to different results. Only a few representatives from each individual facility could participate in the group sessions and we could not verify whether their participation was representative of the opinions expressed by all of their team members. The high participations rates of individual organizations and networks support the generalizability of the results.

Conclusion

This study reports the usefulness of consensus group sessions to reconcile the perceptions of stakeholders about the important dimensions to include in an evaluation of network performance. Since our study is the first to illustrate such a process, repeating similar experiences is required to determine the advantages and drawbacks of different processes to reconcile divergent stakeholder perspectives. Moreover, it remains to be demonstrated how implicating stakeholders from an early stage affects the evaluation process. In the future, the exploration and reconciliation of stakeholder' perspectives about performance evaluation could become an important step in the planning of any performance evaluation process.

Acknowledgements

Preparation of this article was supported by scholarships awarded to the first author by the Canadian Institutes for Health Research, the ANÉIS program, and the Programme de Recherche en Réadaptation et en Intégration Sociale (PRISST).

Reviewers

Edgar Borgenhammar, PhD, MBA, DSI, Professor of Health Services Management, Sweden

Louis Demers, Prof, École nationale d'administration publique [National School of Public Administration], Québec, Canada

Robyn Keast, Associate Professor, School of Management, Faculty of Business and Research Director, Airport Metropolis Project, Faculty of Built Environment and Engineering, Queensland University of Technology, Australia.

References

1. Borgatti S. The network paradigm in organizational research: a review and typology. *Journal of Management* 2003;29(6): 991–1013.
2. Goodwin N, Perri 6, Peck E, Freeman T, Posaner R. Managing across diverse networks of care: lessons from other sectors. London: NCCSDO; 2004. [cited Nov 9 2010]. Available from: <http://www.sdo.nihr.ac.uk/files/adhoc/39-policy-report.pdf>.
3. Kenis P, Provan KG. Towards an exogenous theory of public network performance. *Public Administration* 2009;87(3):440–56.
4. Waldfoegel J. The new wave of service integration. *The Social Service Review* 1997;71(3):463–84.
5. Contandriopoulos A-P, Denis J-L, Touati N, Rodriguez R. Intégration des soins: dimensions et mise en oeuvre [Care integration: dimensions and practice]. *Rupture: Revue Transdisciplinaire en Santé* 2001;8(2):38–52. [in French].
6. Keast R, Brown K, Mandell M. Getting the right mix: unpacking integration meaning and strategies. *International Public Management Journal* 2007;10(1):9–33.
7. Keast R, Mandell MP, Brown K, Woolcock G. Network structures: working differently and changing expectations. *Public Administration Review* 2004;64(3):363–71.
8. Provan KG, Milward HB. A preliminary theory of interorganizational effectiveness: a comparative study of four community health systems. *Administrative Science Quarterly* 1995;40(1):1–33.
9. Fleury M-J. Integrated service networks: the Quebec case. *Health Services Management Research* 2006 Aug;19(3):153–65.
10. Mandell M, Keast R. Evaluating network arrangement. Toward revised performance measure. *Public Performance and Management Review* 2007;30(4):574–97.
11. Mandell MP, Keast R. Evaluating the effectiveness of interorganisational relations through networks: developing a framework for revised performance measure. *Public Management Review* 2008;10(6):715–31.
12. Lamontagne ME, Swaine BR, Lavoie A, Champagne F, Marcotte AC. Perceptions of traumatic brain injury network participants about network performance. *Brain Injury* 2010;24(6):812–22.
13. Leggat S, Leatt P. Cadre d'évaluation du rendement des systèmes de prestation de services de santé intégrée [Performance evaluation of integrated health systems]. *Healthcare Management Forum* 1997;10(1):19–26. [in French].
14. Huxham C. Theorizing collaboration practice. *Public Management Review* 2003;5(3):401–23.
15. Provan KG, Veazie MA, Staten LK, Tufel-Shone NI. The use of network analysis to strengthen community partnership. *Public Administration Review* 2005;65(5):603–13.
16. Huxham C, Vangen S. *Managing to collaborate: the theory and practice of collaborative advantage*. New York: Routledge; 2005.
17. Connolly T, Conlon EJ, Deutch SJ. Organizational effectiveness: a multiple-constituency approach. *Academy of Management Review* 1980;5(2):211–7.
18. Guisset A-L, Sicotte C, Leclercq P, D'Hoore W. Définition de la performance hospitalière: une enquête auprès des divers acteurs stratégiques au sein des hôpitaux [Definition of hospital performance: unease in front of the diversity of strategic actors at the hospitals]. *Sciences Sociales et Santé* 2002;20(2):65–102. [in French].
19. Minvielle E, Sicotte C, Champagne F, Contandriopoulos AP, Jeantet M, Preaubert N, et al. Hospital performance: competing or shared values? *Health Policy* 2008;87(1):8–19.
20. Tregunno D, Ross Baker G, Barnsley J, Murray M. Competing values of emergency department performance: balancing multiple stakeholder perspectives. *Health Services Research* 2004;39(4 Pt 1):771–91.
21. Finucane PM, Barron SR, Davies HA, Hadfield-Jones RS, Kaigas TM. Towards an acceptance of performance assessment. *Medical Education* 2002;36(10):959–64.
22. Gagliardi A, Lemieux-Charles L, Brown A, Sullivan T, Goel V. Stakeholder preferences for cancer care performance indicators. *International Journal of Health Care Quality Assurance* 2008;21(2):175–89.
23. Sicotte C, Champagne F, Contandriopoulos AP, Barnsley J, Beland F, Leggat SG, et al. A conceptual framework for the analysis of health care organizations' performance. *Health Service Management Research* 1998;11(1):24–41. discussion-8.
24. Leutz WN. Five laws for integrating medical and social services: lessons from the United States and the United Kingdom. *The Milbank Quarterly* 1999;77(1):77–110. iv–v.
25. Zinn J, Zalokowski A, Hunter L. Identifying indicators of laboratory management performance: a multiple constituency approach. *Health Care Management Review* 2001;26(1):40–53.
26. Dijk JAGMV. Delphi Questionnaires versus groups interviews. A comparison case. *Technological Forecasting and Social Change* 1990;(37):293–304.
27. Bryson JM, Crosby BC, Stone MM. The design and implementation of cross-sector collaborations: propositions from the literature. *Public Administration Review* 2006;66(s1):44–55.
28. Gervais M, Pépin G. TRIAGE: a new group technique gaining recognition in evaluation. *Evaluation Journal of Australasia* 2002;2(2):45–9.
29. Swaine B, Dutil E, Demers L, Gervais M. Evaluating Clients' perception of the quality of head in jury rehabilitation services: development and validation of a questionnaire. *Brain Injury* 2000;7:575–87.