

Volume 12, 18 September 2012

Publisher: Igitur publishing

URL: <http://www.ijic.org>

URN:NBN:NL:UI:10-1-113112 / ijic2012-137

Copyright: 

Submitted: 31 January 2012, revised 26 June 2012, accepted 26 June 2012

---

Research and theory

## Measuring the degree of integration for an integrated service network

*Chenglin Ye, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton ON, L8S 4L8 Canada*

*Gina Browne, School of Nursing, McMaster University, Hamilton ON, L8S 4L8 Canada*

*Valerie S. Grdisa, Healthcare Advisory, Management Consulting, KPMG Canada, Toronto ON, Canada*

*Joseph Beyene, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton ON, L8S 4L8 Canada*

*Lehana Thabane, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton ON, L8S 4L8 Canada*

*Correspondence to: Lehana Thabane, Biostatistics Unit, Father Sean O'Sullivan Research Centre, Room H325, St. Joseph's Healthcare Hamilton, 50 Charlton Avenue East, Hamilton ON, L8N 4A6 Canada, Phone: +1/905-522-1155 ext. 33720, Fax: +1/905-308-7212, E-mail: [thabanl@mcmaster.ca](mailto:thabanl@mcmaster.ca)*

---

### Abstract

**Background:** Integration involves the coordination of services provided by autonomous agencies and improves the organization and delivery of multiple services for target patients. Current measures generally do not distinguish between agencies' perception and expectation. We propose a method for quantifying the agencies' service integration. Using the data from the Children's Treatment Network (CTN), we aimed to measure the degree of integration for the CTN agencies in York and Simcoe.

**Theory and methods:** We quantified the integration by the agreement between perceived and expected levels of involvement and calculated four scores from different perspectives for each agency. We used the average score to measure the global network integration and examined the sensitivity of the global score.

**Results:** Most agencies' integration scores were <65%. As measured by the agreement between every other agency's perception and expectation, the overall integration of CTN in Simcoe and York was 44% (95% CI: 39%–49%) and 52% (95% CI: 48%–56%), respectively. The sensitivity analysis showed that the global scores were robust.

**Conclusion:** Our method extends existing measures of integration and possesses a good extent of validity. We can also apply the method in monitoring improvement and linking integration with other outcomes.

### Keywords

integration measure, perception, expectation, collaboration agreement

---

## Background

Spending on healthcare in Canada continues to outpace government revenue and economic growth [1]. In 2010, the total Canadian healthcare expenditures were approximately \$191.6 billion [2]. Fraser Institute, a Canadian think-tank, forecasts that if the recent trend continues, six of 10 provincial governments will spend more than half of total revenue on healthcare by the year 2020 [3]. Many factors contribute to escalating healthcare costs that include: development of new drugs, availability of expensive health services, and an increase in co-morbid chronic conditions in our aging populations [4–6]. The sustainability of the existing Canadian universal care system is a growing concern [7].

In general, patients with chronic illness or functional limitations are the major consumers of community health services [8, 9]. For example, in addition to taking psychopharmacologic drugs, patients with mental illness would need psychiatric counseling, rehabilitative therapy, and other health and social services to help them and their families. Usually, such primary care services are delivered by community health centres—which are operated autonomously [10]. Patients often receive help for a single problem at a time and endure duplicative processes to receive multiple services. With limited resources, we can hardly meet the total needs of a population by providing health services in an isolated and fragmented fashion. As a result, patients may experience prolonged wait-times and courses of treatment.

Over the last decade, integration has been advocated as a viable strategy for improving the organization and delivery of health services. A majority of the literature on integrating services discusses the integration of health services on primary care, hospital care, or mental health care [8–11]. Often, the integration of health services is framed under the notion of ‘continuity of care’ [12]. A prominent hypothesis is that integration of health services leads to higher quality of care at a lower cost and maintains or improves patients’ health and satisfaction [13, 14]. In this paper, we adopt the definition introduced by Browne et al., who define integration as the coordination of a comprehensive spectrum of services (e.g., health, education, community and social services) provided by multiple agencies [10]. We use providers and agencies interchangeably to represent the publicly funded organizations that deliver community health services.

The integrated approach of delivering services has potential advantages. Patients will receive comprehensive care and are likely to gain a much better outcome. Another advantage is that service providers can

improve their caseload by coordinating with other providers and reducing duplication, and thus, better meet the overall needs of a population. There are different ways to initiating integration within a community. One is by legislative change, such as the Health Action Zones in the United Kingdom whereby a whole geographic area receives the total funding for public and health services. The implementation of Local Health Integration Networks (LHINs) in Ontario, Canada is another example where local region receives funds for acute, community, and long-term care services. Regional health authorities develop service agreements with performance indicators. Savings are kept within the region and used for other priorities. The other approach begins at a local level whereby agencies collaborate together to serve more clients without increasing funding allocation. The local community develops own collaboration in service delivery. The Children’s Treatment Network (CTN) in Ontario, Canada is an example of the third approach, for serving the children with complex health problems.

Studies have shown that integrating local services improves patient outcomes, such as reducing functional decline in the elderly [15], preventing avoidable hospitalizations [16], and minimizing the risk of developing diabetes-related complications [17]. One sentence summarizes the current evidence: healthcare providers can achieve target outcomes more easily with less investment by coordinating available, necessary, and preferable human services to patients [18]. From a societal perspective, proactive and comprehensive services are more effective and less expensive because giving people what they need in a coordinated fashion results in a reduced use of other services [19]. The emphasis on human services is the result of accumulating evidence that the factors (besides genetic predisposition) determining health are social, environmental, educational, and personal in nature. However, the extent to which those services are integrated for addressing patients’ health needs is often unknown and rarely measured.

Without quantifying the degree of integration, the effect of collaborative effort can be hardly identified. Without measuring and monitoring the actual integration, it is difficult for network planners to make decisions and confirm successful implementation [20]. Thus, we need a valid method for measuring the degree to which a network has achieved in integrating services. A recent systematic review [13] identified 18 quantitative measures of integration. None of those measures separate agencies’ current perception of involvement from their expectation. Since agencies would differ in the level of involvement necessary to meet patients’ needs, a rational measure should reflect the agreement between perceived and expected levels of involvement.

In marketing, researchers have used the gap between the expected performance and the perceived experience as an objective measure of customer satisfaction [21]. Current measures generally do not distinguish between agencies' perception and expectation on integration although separate measures offer an appropriate and reliable way to identify the 'gap' in integration [22]. Browne et al. develop the Human Service Integration Measure (HSIM) scale [10] that includes separate measures of observed and expected integrations. Another key limitation in existing measures is the common assumption that a higher level of involvement is always a better integration. A high level of input by an agency is not necessarily an agreed-upon involvement by all agencies [10, 13, 23]. A measure of integration needs to account for the expectations of all underlying agencies in their pursuit of a common goal to improve patient outcomes [24–26]. In this paper, we introduce a new method to measure the degree of service integration among agencies. Our proposed method has unique strengths by: quantifying the gap between perceived and expected levels of involvement within a network; creating a relational score by the agreement among all agencies; measuring the integration from four different perspectives; and creating integration scores for individual agencies and global integration scores for the entire network as a whole. In this paper, we use the words 'perceived', 'observed', and 'actual' interchangeably to mean the current level of involvement perceived by agencies. We use the words 'expected' or 'optimal' to mean the level of involvement expected by agencies. Using the data from the Children's Treatment Network (CTN), our objectives are to: 1) measure the degree of integration among agencies of CTN; 2) calculate global integration scores of CTN; and 3) assess the sensitivity of the global integration score based on different approaches for estimating the global integration score.

## Theory and methods

### Study design and study population

The data were drawn from a cross-sectional study that evaluated the integration of CTN in 2006, one year after its inception. The study measured the degree of integration among agencies of CTN in Simcoe and York separately.

In Simcoe County and York Region of Ontario, Canada, families with severely disabled child(ren) had limited access to specialized treatments and had to travel outside their region to receive services. They used services in a disjointed and self-directed manner and often received suboptimal outcomes. In response

to these problems, Simcoe County and York Region launched the CTN in 2005. Targeting the children with complex health problems, local service agencies in CTN collaborated together to deliver comprehensive therapeutic and psychosocial services to the children and their families. This innovative and proactive model integrated existing service agencies—including health, recreational, educational, social, mental health, and community resources ([www.ctn-simcoeyork.ca](http://www.ctn-simcoeyork.ca)). Each family served by CTN had a unique team of service providers for a long-term basis. The interdisciplinary team provided a single point of contact, health assessment, service coordination, and a comprehensive plan of care for the children.

### Characteristics of the CTN agencies

At the time of measurement, there were 27 and 36 agencies in Simcoe and York, respectively. The CTN represented health, educational, social, justice, recreational, and cultural sectors. There were different types of agencies in each sector that included: early years, Healthy Babies, adolescent support, rehabilitation, home care, social assistance, child protection, mental health, recreation, leisure services, etc.

### The Human Service Integration Measure Scale

To quantify the level of partnership involvement, we used the latest version of the Human Service Integration Measure (HSIM) scale, a 5-point ordinal scale developed by Browne et al., to validate their integration framework [10]. Representatives of agencies can fill out the measures by web-form, phone, or in-person interview. An interviewer will ask each agency to rate its current and expected levels of involvement with the other agencies within the network. An example of the measure is provided in Figure 1. For agencies that have more than one representative, the interviewer will average the ratings from all representatives and round it to the nearest integer.

### Organizing the data for calculating integration

We have developed two square matrices to organize the responses of perceived involvement and expected involvement, respectively. Each column in the matrix contains an agency's ratings on every other agency (i.e., self-ratings) and each row contains other agencies' ratings on the same agency (i.e., group-ratings). An illustrative example is provided in Figure 2. In the matrix of perceived involvement, the first column contains the level of involvement that Agency A has

To what extent <b>are you</b> involved with the following services?	Agency/			To what extent <b>should you</b> be involved with the following services?
Rating scale	Rate (0-4)	Do not rate your own service	Rate (0-4)	Rating scale
<p><b>0=No awareness:</b> Your agency <b>is not</b> aware of the other service</p> <p><b>1=Awareness:</b> You <b>have</b> knowledge of the other service although no effort is taken to organize activities according to any principles except those that conform to individual agency missions.</p> <p><b>2=Communication:</b> You and the other service <b>have</b> an active program of communication and information sharing.</p> <p><b>3=Cooperation:</b> You and the other service <b>each use</b> your knowledge of the other's service to guide and modify your own service planning in order to obtain a better set of links between services.</p> <p><b>4=Collaboration:</b> You and the other service <b>jointly</b> plan the offering of service and actively modify service activity based on advice and input from mutual discussions.</p>		Agency A		<p><b>0=No awareness:</b> Your agency <b>should not</b> be aware of the other service</p> <p><b>1=Awareness:</b> You <b>should</b> have knowledge of the other service although no effort is taken to organize activities according to any principles except those that conform to individual agency missions.</p> <p><b>2=Communication:</b> You and the other service <b>should</b> have an active program of communication and information sharing.</p> <p><b>3=Cooperation:</b> You and the other service <b>should</b> each use your knowledge of the other's service to guide and modify your own service planning in order to obtain a better set of links between services.</p> <p><b>4=Collaboration:</b> You and the other service <b>should</b> jointly plan the offering of service and actively modify service activity based on advice and input from mutual discussions.</p>
		Agency B		
		Agency C		
		Agency D		
		Agency E		
		Agency F		
		Agency G		
		Agency H		
		Agency I		
		Agency J		

Figure 1. The Human Service Integration Measure Scale.

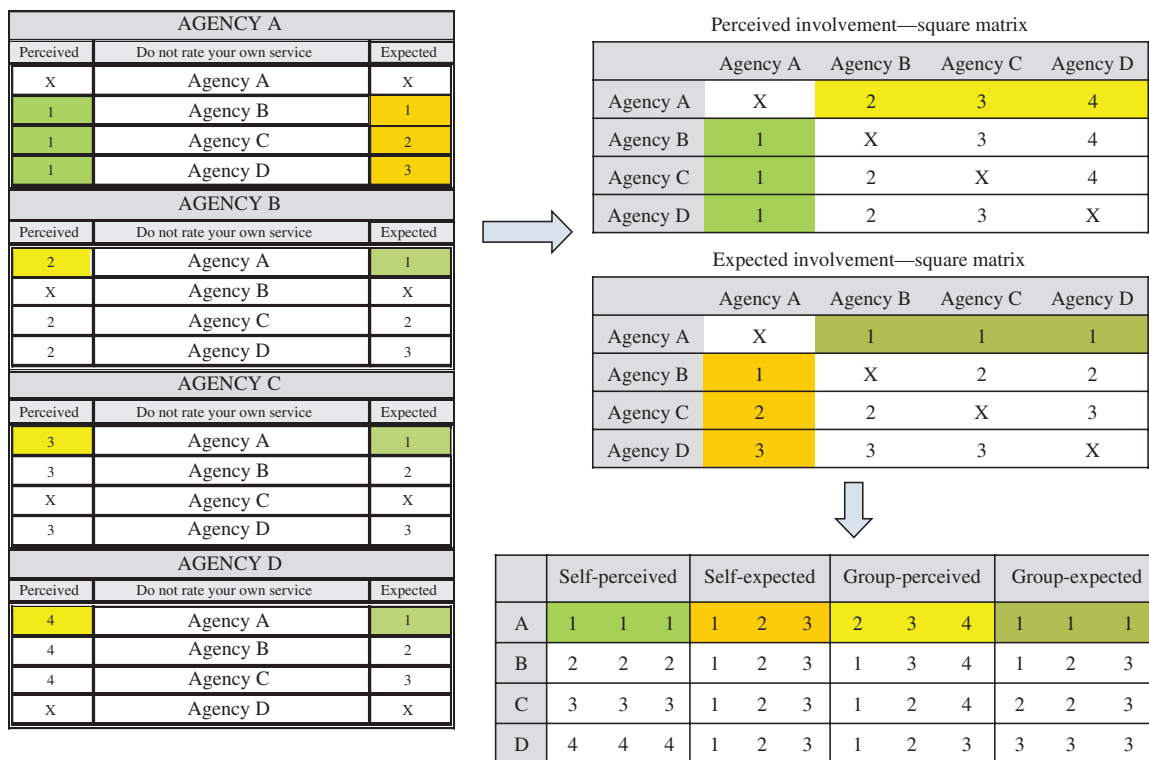


Figure 2. An example for organizing the data.

perceived with every other agency (e.g., 1, 1, and 1) and the first row contains the level of involvement that other agencies have perceived with Agency A (e.g., 2, 3 and 4). Similarly, in the matrix of expected involvement, the first column contains the level of involvement that Agency A has expected with every other agency (e.g., 1, 2, and 3) and the first row contains the level of involvement that other agencies have expected with Agency A (e.g., 1, 1, and 1). The ratings for other agencies are organized in the same way. Every agency has four types of ratings, namely, self-perceived, self-expected, group-perceived, and group-expected ratings.

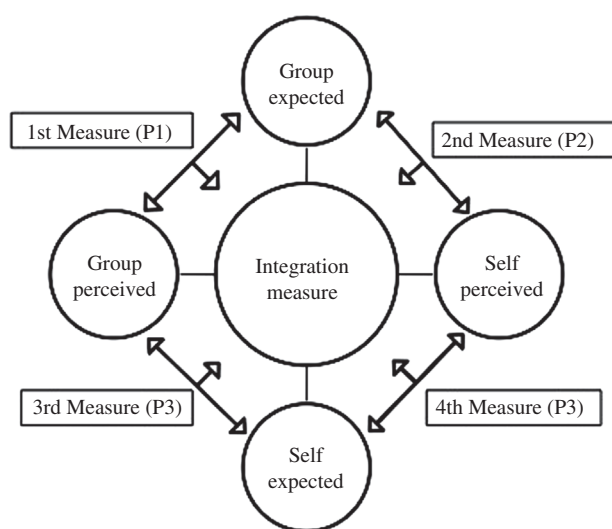
### Calculating the agency integration scores

We measured an agency’s integration score by the agreement between perceived and expected involvements with other agencies and defined the agreement as the percentage of pairs of agreed perceived and expected ratings. Thus, our integration score was an agreed-upon score on the level of involvement among agencies. For example, if there were 80% of agencies whose group-perceived scores on Agency X were same as their group-expected scores, then

the corresponding degree of integration for Agency X would be 80%. Our integration framework measured the agreement from four perspectives shown in Figure 3, including the agreement between: the group-perceived and group-expected involvements (P1), the self-perceived and group-expected involvements (P2), the group-perceived and self-expected involvements (P3), and the self-perceived and self-expected involvements (P4).

### Calculating the global integration score

We estimated the global integration score of a network by the average integration score. As shown in Figure 4, the graded area represented the global integration score of a network based on scores from all the network agencies. The blank area on the diagram indicates the gap in the degree of global network integration. The total area is the sum of the graded and the blank areas of the diagram which equals 1 (i.e., 100% integration). Essentially, the graded area represents the average integration score. We calculated corresponding global integration score with 95% confidence interval (CI) as the graded area based on P1, P2, P3, and P4 integration scores.



Integration score	Agreement between...	It measures...
P1	Group-perceived rating and group-expected rating	The degree of agreement between an agency’s involvement perceived by others and its involvement expected by others.
P2	Self-perceived rating and group-expected rating	The degree of agreement between an agency’s self-perceived involvement and its involvement expected by others.
P3	Group-perceived rating and self-expected rating	The degree of agreement between an agency’s involvement perceived by others and the involvement expected itself.
P4	Self-perceived rating and self-expected rating	The degree of agreement between an agency’s self-perceived involvement and the involvement expected itself.

Figure 3. The framework of measuring integration.

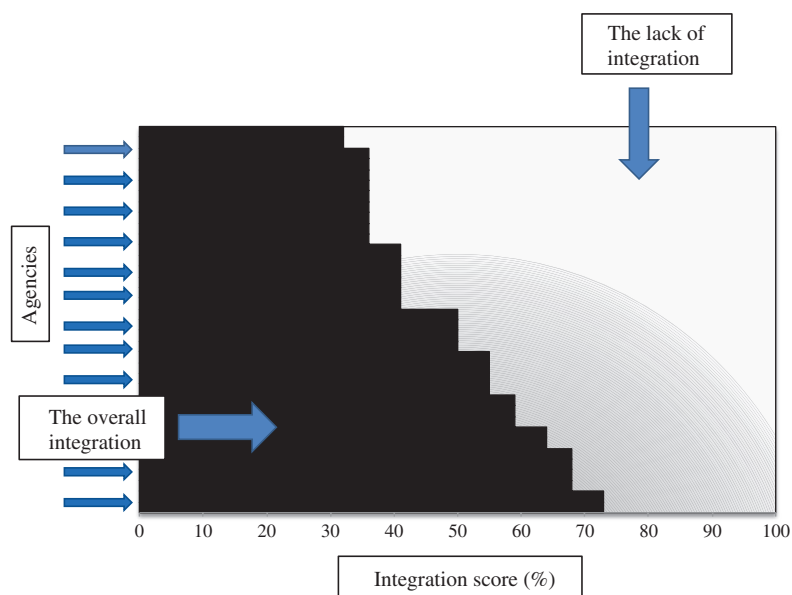


Figure 4. An example of the global integration.

## Assessing the sensitivity of global integration score

We conducted a sensitivity analysis by comparing with other methods for estimating the global integration score: the weighted-average method and the bootstrap method. In the weighted-average method, we attached a different weight ( $w$ ) to each agency based on the variance of its integration score. This method adjusted for the precision of an estimated integration score. An agency with a more precise integration score contributed more to the global score for the whole network. For the bootstrapping method, we used three different bootstrap procedures: the standard, the balanced, and the Bayesian procedures. Bootstrap is a common resampling method for improving estimation and confidence intervals of an unknown parameter [27–29]. Different procedure requires a different resampling algorithm and thus, estimates the sampling error differently. The standard bootstrap can produce a bias-corrected estimate [29] that largely reduces the potential bias arising in estimation. The balanced bootstrap is similar to the standard procedure but bootstrap samples are balanced. Compared with the standard bootstrap, the balanced bootstrap generally improves the efficiency of simulation [30]. The Bayesian bootstrap uses a different algorithm and approximates a posterior distribution of the global score instead of a sampling distribution [31]. For both the standard and balanced bootstrap procedures, we computed the 95% bias-corrected and accelerated CI that adjusted both bias and skewness in bootstrap sampling [28]. For the

Bayesian bootstrap procedure, we computed the 95% credibility interval (CrI) instead. The details of calculations and procedures were provided in the Appendix. We performed all statistical analyses in the software package R version 2.12.1.

## Results

### The integration scores for agencies of CTN

We summarized the integration scores for CTN agencies in Table 1. For confidentiality, we kept agencies anonymous and labeled them by Arabic numbers. The response rate was 89% in Simcoe and 64% in York. All agencies (i.e., both respondents and non-respondents) received a P1 score. However, only respondents received P2, P3, and P4 scores because those calculations required the agencies' ratings on other agencies.

In Simcoe, P1, P2, P3, and P4 integration scores varied from 25% to 82%, 4% to 83%, 13% to 83%, and 4% to 96%, respectively. In York, P1, P2, P3, and P4 integration scores varied from 27% to 73%, 27% to 81%, 9% to 82%, and 23% to 97%, respectively.

### The global integration scores for the CTN

Global integration scores of CTN were summarized in Table 2. In Simcoe, P1, P2, P3, and P4 global integration

**Table 1.** Integration scores for the agencies of the Children’s Treatment Network (CTN)

Service agency	Simcoe (n=27)				York (n=36)			
	P1	P2	P3	P4	P1	P2	P3	P4
1	30%	39%	26%	12%	41%	62%	14%	24%
2	38%	–	–	–	68%	59%	59%	46%
3	43%	30%	35%	31%	35%	–	–	–
4	43%	26%	83%	88%	41%	50%	27%	43%
5	39%	43%	26%	23%	52%	–	–	–
6	61%	65%	35%	54%	68%	32%	64%	26%
7	35%	35%	65%	38%	52%	–	–	–
8	25%	–	–	–	61%	–	–	–
9	25%	–	–	–	64%	36%	45%	31%
10	48%	26%	48%	50%	73%	68%	82%	79%
11	39%	35%	22%	12%	57%	–	–	–
12	43%	30%	43%	31%	59%	50%	64%	74%
13	26%	35%	22%	19%	64%	59%	68%	49%
14	82%	55%	82%	88%	50%	64%	59%	74%
15	57%	39%	65%	38%	43%	–	–	–
16	52%	45%	41%	12%	43%	–	–	–
17	26%	43%	61%	96%	68%	55%	82%	40%
18	48%	61%	43%	85%	55%	45%	73%	49%
19	48%	48%	57%	65%	64%	–	–	–
20	48%	4%	57%	4%	41%	55%	50%	57%
21	48%	83%	13%	58%	32%	36%	9%	24%
22	48%	48%	52%	42%	36%	82%	41%	97%
23	43%	74%	13%	50%	41%	48%	62%	68%
24	52%	35%	39%	15%	43%	–	–	–
25	52%	70%	22%	19%	59%	73%	45%	49%
26	48%	35%	30%	69%	62%	71%	52%	51%
27	39%	30%	57%	50%	27%	–	–	–
28	–	–	–	–	57%	–	–	–
29	–	–	–	–	55%	27%	50%	23%
30	–	–	–	–	43%	–	–	–
31	–	–	–	–	36%	50%	27%	31%
32	–	–	–	–	68%	41%	82%	51%
33	–	–	–	–	73%	68%	77%	77%
34	–	–	–	–	57%	–	–	–
35	–	–	–	–	41%	41%	64%	66%
36	–	–	–	–	48%	76%	48%	69%

Percentages were rounded up to integer; the agencies were listed in a consecutive order and the same number **did not** refer to the same agency; CI=confidence interval.

P1=the degree of agreement between an agency’s involvement perceived by others and its involvement expected by others.

P2=the degree of agreement between an agency’s self-perceived involvement and its involvement expected by others.

P3=the degree of agreement between an agency’s involvement perceived by others and the involvement expected itself.

P4=the degree of agreement between an agency’s self-perceived involvement and the involvement expected itself.

scores with 95% CI were 44% (39%, 49%), 43% (36%, 51%), 43% (35%, 52%), and 44% (32%, 55%), respectively. In York, they were 52% (48%, 56%), 54% (48%, 61%), 54% (45%, 63%), and 52% (43%, 61%), respectively. The global integration of CTN in York was generally higher than that in Simcoe.

### Assessing the global integration score

The global integration scores calculated by different approaches were similar. We performed all bootstrap procedures by simulating 500, 1000, 5000, 10,000, and 40,000 bootstrap samples. Although, increasing the number of simulations reduced the random sampling error caused by the bootstrap procedure itself,

the results only differed in the third decimal place. Thus, we only reported the results by simulating 40,000 bootstrap samples. The weighted-average approach provided a slightly different estimate in some cases and the narrowest 95% CI (Figures 5 and 6). Other researchers also reported a narrower confidence interval when using a weighted approach [32]. Still, the 95% CIs covered all scores by different methods. The only exception was the P4 global score measuring the overall agreement between self-perceived and self-expected involvements in York, where the weighted-average method provided a significantly larger estimate than other methods. The global scores estimated by different bootstrap procedures were identical to the standard one. This showed that the average integration

**Table 2.** Global integration scores estimated by different methods

Region	Method	P1	95% CI*	P2	95% CI*	P3	95% CI*	P4	95% CI*
Simcoe	Sample mean	44%	(39%, 49%)	43%	(36%, 51%)	43%	(35%, 52%)	44%	(32%, 55%)
	<b>Other methods</b>								
	Weighted average	44%	(40%, 48%)	38%	(34%, 42%)	41%	(37%, 45%)	45%	(42%, 48%)
	Standard bootstrap	44%	(39%, 49%)	43%	(37%, 50%)	43%	(35%, 51%)	44%	(33%, 55%)
	Balanced bootstrap	44%	(39%, 49%)	43%	(36%, 50%)	43%	(35%, 51%)	44%	(33%, 55%)
Bayesian bootstrap	44%	(40%, 49%)	43%	(37%, 50%)	43%	(36%, 51%)	44%	(34%, 55%)	
York	Sample mean	52%	(48%, 56%)	54%	(48%, 61%)	54%	(45%, 63%)	52%	(43%, 61%)
	<b>Other methods</b>								
	Weighted average	52%	(49%, 56%)	55%	(51%, 60%)	52%	(48%, 56%)	62%	(59%, 65%)
	Standard bootstrap	52%	(48%, 56%)	54%	(48%, 60%)	54%	(46%, 62%)	52%	(44%, 60%)
	Balanced bootstrap	52%	(48%, 56%)	54%	(48%, 60%)	54%	(46%, 62%)	52%	(44%, 60%)
Bayesian bootstrap	52%	(48%, 56%)	54%	(48%, 60%)	54%	(46%, 62%)	52%	(44%, 60%)	

\*CI=confidence interval, a credibility interval was calculated instead in the Bayesian bootstrap method; n=number of respondents; bootstrap estimates were obtained by simulating 40,000 bootstrap samples.

P1=the degree of agreement between an agency's involvement perceived by others and its involvement expected by others.

P2=the degree of agreement between an agency's self-perceived involvement and its involvement expected by others.

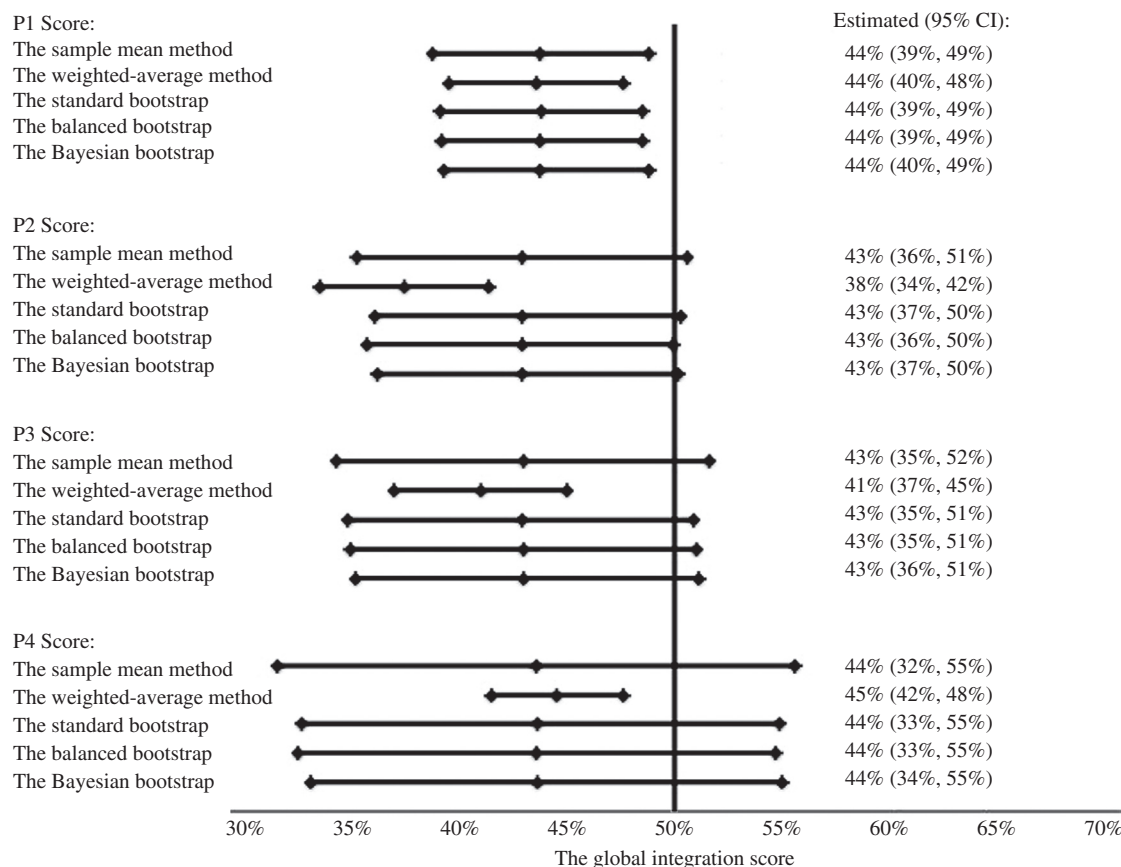
P3=the degree of agreement between an agency's involvement perceived by others and the involvement expected itself.

P4=the degree of agreement between an agency's self-perceived involvement and the involvement expected itself!

score was a simple and reliable estimate of the global score. The findings were consistent with the fact that a sample mean was an unbiased estimate of the population mean. The 95% CIs in bootstrap methods were slightly more precise. Overall, the sample mean was a robust estimate of the global integration score.

## Discussion

We have developed a method for quantifying the degree of integration for agencies in an integrated service network. Using this method, managers could identify the current gap in service integration. We applied



**Figure 5.** Global integration scores of CTN Simcoe.



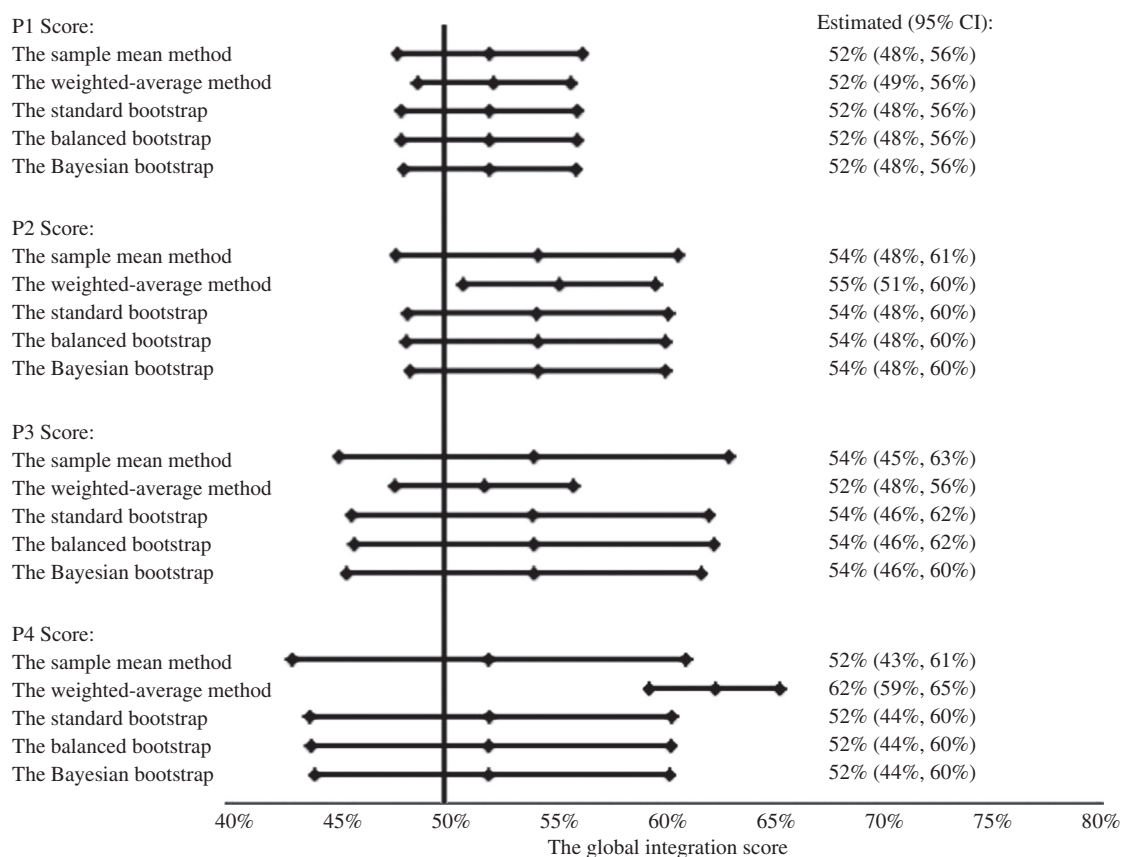


Figure 6. Global integration scores of CTN York.

the method in measuring the CTN agencies. Non-respondent agencies had lower P1 integration score, which indicated a poorer integration as perceived and expected by the group. For some agencies, their scores (i.e., P1, P2, P3, and P4 scores) varied largely across different perspectives of integration. As shown in Figures 7 and 8, the spider plots were helpful for examining the gap of integration by different views. When an agency had 100% integration, the plot would show a complete ‘diamond’. Any defect on the ‘diamond’ would indicate imperfect integration from some perspective. For example, in Simcoe, Agency 4 had much lower P1 and P2 scores than its other scores: below 45% vs. above 80%. This suggested that the level of involvement perceived by all agencies including Agency 4 itself did not meet other agencies’ expectation. This was an indication that the group might have a wrong expectation on the level of involvement required for Agency 4. Agency 14 had a much lower P2 score than other scores: 55% vs. above 80%. The level of involvement perceived by Agency 14 met its own expectation but not others. This was an indication that Agency 14 might have an improper perception on the level of involvement it was contributing to the group. Agencies 17 had much lower P1, P2, and P3 scores than its P4

score: below 62% vs. above 85%. This showed that the level of involvement perceived by Agency 17 only met its own expectation but not others, and was different from others’ perception too. Agency 21 had much lower P1, P3, and P4 scores than its P2 score: below 59% vs. above 82%. Although Agency 21’s perception on its current involvement met others’ expectation, it was different from what others had perceived. Our findings helped CTN managers to diagnose deeper problems of integration and potential barriers in integrating multiple services. By measuring the service integration over time, we could also evaluate the improvement in agency working relationships and promote further dialogue in achieving better integration.

Our measurement had some limitations. First, we only focused on measuring the degree of collaborative involvement in service integration. There were other types of integration [33], e.g., the functional integration, which we did not measure. Second, the measurement only captured the integration among the planning group. We acknowledged that integration achieved at the planning table did not always reflect the degree of integration in real practice, for example, among the frontline teams of workers. Third, there was

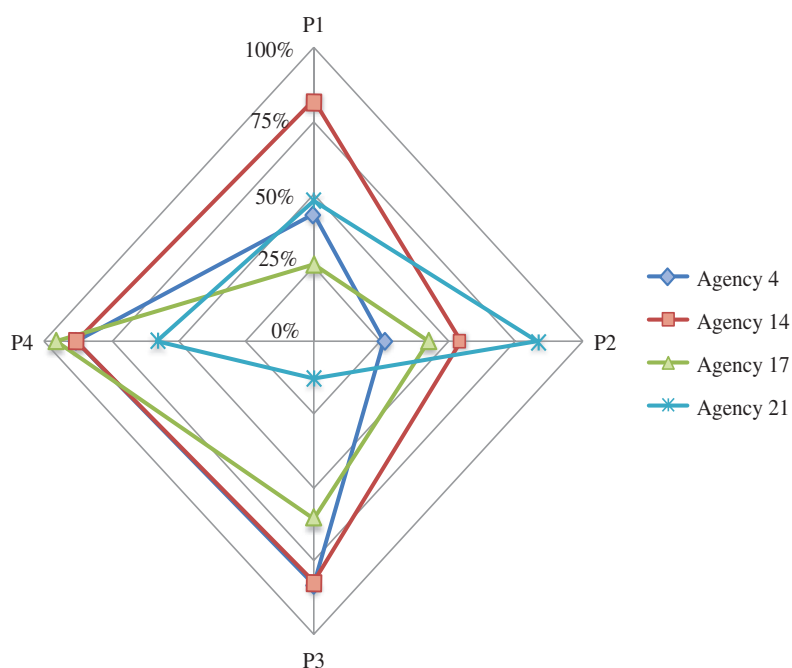


Figure 7. The spider plot for comparing the 4 integration scores for agencies in Simcoe.

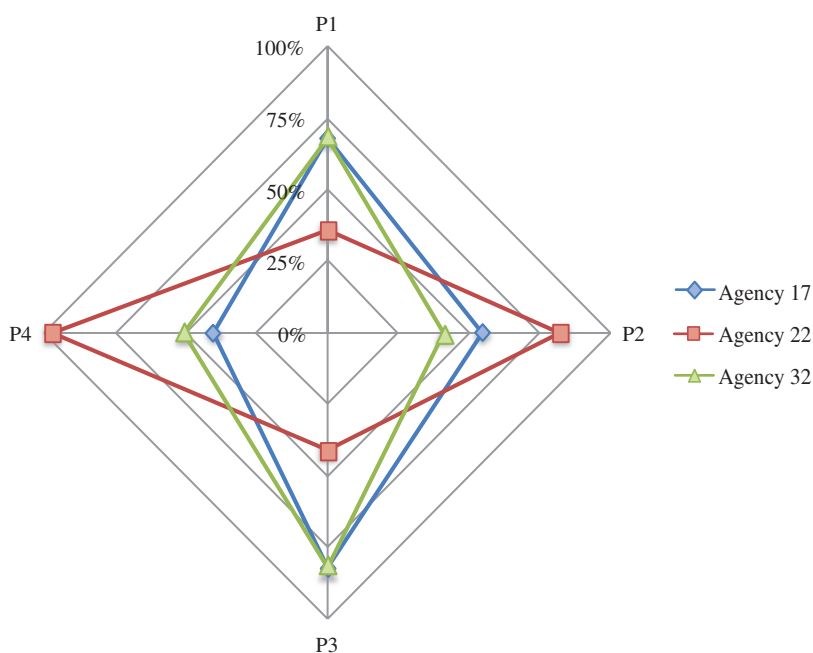


Figure 8. The spider plot for comparing the 4 integration scores for agencies in York.

potential respondent bias because filling out the measure by representative(s) is subject to proxy reporting bias. Representatives might not give the same information that others from the same agency would give. Halo effect and end-aversion were two other potential sources of bias in our results [34]. Finally, we were not

able to evaluate the impact of non-respondents on the results.

Integration models generally require some formal mechanism, such as networks or committees of local agencies, to plan, organize, and deliver multiple

services together. There are many barriers to integration because current health, educational, social, rehabilitative services, etc. are funded independently. Relationship, politics, communication, process, structure, and conflict are common problems for the failure of integration [35–38]. The gains from integration are often difficult to sustain and we need a tool to measure them. Despite the limitations, our method provides a valid way to conceptualize and quantify the service integration among agencies. During the time of measurement, CTN was undergoing initial planning to link resources and support, organize services, and create a new governance model. This partly explained the low degree of integration that CTN agencies had achieved. By quantifying the state of integration, our measurement helped CTN agencies visualize their agreement in the process of integration and generated important discussions for their next stage of planning. The unique value of our tool is that it provides a relative score on the degree to which the actual involvement agrees with the expected involvement. In this regard, our method greatly extends the HSIM by bridging the perceived and the expected integrations. Compared with the Ahgren and Axelsson's method [23], our approach does not adopt an assumption that higher levels of integration are better. Another difference was that the Ahgren and Axelsson's method uses specific criteria to define levels of integration (e.g., referrals, guidelines, chains of care, network managers, and pooled resources). Although this might be a more objective approach, the measure could not distinguish the differences in the degree of involvement that should be present in a well-functioning network. In addition, our method produces simple and reliable global integration scores that quantify the integration of an entire network as a whole, which is hardly addressed in the current literature. Our method possesses a good content relevance and coverage on measuring integration by including: the spectrum of services, the number of providers, the score of integration, and the perspective of agencies. Our method can also differentiate the degree of integration between similar networks, as shown in case of the CTN Simcoe and York.

We have used the Partnership Self-Assessment Tool [39, 40] to examine the association between our integration scores and components of a collaborative process—synergy, leadership, efficiency, administration and management, resources, decision-making, benefits, drawbacks, and satisfaction. Our results showed that synergy was strongly associated with integration [41]. Other components including leadership, administration, decision-making, and satisfaction were also associated with integration. The findings demonstrated some extent of convergent validity because our integration measure was related to other variables of the

same construct to which it should be related [34]. In an on-going analysis, we are examining the linkage between the degree of integration and network capacity that includes the average wait-time and the caseload. For future studies, we can apply the measurement in other similar service networks or in the same network for continuous evaluation. By repeating the measurement, investigators can determine integration patterns over time and examine the connection between integration and network outcomes longitudinally. We have found a potential influence of provider team integration on the quality of life of children with complex needs [42]. Future studies can also examine the similar relationship between different patient outcomes and the degree of integration using our measure.

## Concluding remarks

In this paper, we introduce a method for measuring the degree of integration for agencies in an integrated service network. Using the method, we measured the integration of the CTN agencies in Simcoe and York. In CTN Simcoe, agencies' P1, P2, P3, and P4 integration scores varied from 25% to 82%, 4% to 83%, 13% to 83%, and 4% to 96%, respectively. In CTN York, agencies' P1, P2, P3, and P4 integration scores varied from 27% to 73%, 27% to 81%, 9% to 82%, and 23% to 97%, respectively. Most agencies had a score of <65%, a relatively low level of integration. The results revealed existing problems in integrating CTN services. As measured by the agreement between every other agency's perception and expectation, the overall integration of CTN in Simcoe and York was 44% (95% CI: 39%–49%) and 52% (95% CI: 48%–56%), respectively. The sensitivity analysis showed that the average integration score was a reliable and robust estimate of the global integration score. The measurement provided timely information for decision-making and improving the integration of CTN. The key implication is that every integrated service network needs a valid measurement to evaluate whether or not the collaborative process has been implemented as planned. Measuring the service integration should be the first step in this evaluation. However, quantitative methods available for measuring integration have been scarce in the literature. Our method greatly extends the existing measures of integration by quantifying the agreement between agencies' perceived and expected levels of involvement. Our approach is unique such that we quantify the integration from four different perspectives to identify deeper problems in integrating multiple services provided by autonomous agencies. We showed that the proposed method possessed a good extent of validity and could be applied in measuring other integrations in a similar setting.

## Authors' contributions

CY conceived the study, designed the integration scores, performed all analyses, interpreted the results, and drafted and revised the manuscript. GB advised on important intellectual content and revised the manuscript. VSG revised the manuscript and contributed to the interpretation of results. JB revised the manuscript and contributed to the statistical analysis of the study. LT contributed to the statistical design, interpretation of results and revision of the manuscript. All authors have read and approved the final manuscript.

## Acknowledgements

We thank Health and Social Service Utilization Research Unit (HSSURU) at McMaster University for providing the data. Chenglin Ye is supported in part

by funding from the Father Sean O'Sullivan Research Center (FSORC) Studentship award, the Canadian Institute of Health Research (CIHR) Training award in Bridging Scientific Domains for Drug Safety and Effectiveness, and the Canadian Network and Centre for Trials Internationally (CANNeCTIN) programme.

## Reviewers

**Reynaldo Holder**, MD, Regional Advisor—Hospitals and Integrated Health Care Delivery Systems, Area of Health Systems based on Primary Health Care, PAHO/WHO, Washington DC, USA.

**Nick Goodwin**, Senior Fellow, The King's Fund, London, UK.

**Daniel Salhani**, PhD, Associate Professor, School of Social Work, University of British Columbia, Okanagan Campus, Canada.

## References

1. Canadian Institute for Health Information. National health expenditure trends 1975 to 2010. Canada: CIHI; 2010 Oct. [cited 21 May 2012]. Available from: [https://secure.cihi.ca/free\\_products/nhex\\_trends\\_report\\_2011\\_en.pdf](https://secure.cihi.ca/free_products/nhex_trends_report_2011_en.pdf).
2. Canadian Institute for Health Information. Health care in Canada 2010 Canada: CIHI; 2010 Dec. [cited 21 May 2012]. Available from: <https://secure.cihi.ca/estore/productSeries.htm?pc=PCC64>.
3. Skinner BJ, Rovere M. Paying more, getting less: measuring the sustainability of government health spending in Canada 2009 Report. Canada: Fraser Institute; 2009 Nov. [cited 21 May 2012]. Available from: <http://www.ihe.ca/publications/health-db/geo/448/>.
4. The Conference Board of Canada. Understanding health care cost drivers and escalators Ottawa; 2004 Mar. [cited 21 May 2012]. Available from: <http://www.health.alberta.ca/documents/Health-Costs-Drivers-CBC-2004.pdf>.
5. Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. *The Lancet* 2009;374(9696):1196–208.
6. Canadian Health Services Research Foundation. Myth: Canada's system of healthcare financing is unsustainable Ottawa: CHSRF; 2007 Jan. [cited 21 May 2012]. Available from: <http://www.chsrf.ca/PublicationsAndResources/Mythbusters/ArticleView/07-12-01/bb0d6092-9ec8-45c0-a38d-de696b7b25c9.aspx>.
7. Stuart N, Adams J. The sustainability of Canada's healthcare system: a framework for advancing the debate. *Healthcare Quarterly* 2007;10(2):96–102.
8. Denton FT, Spencer BG. Chronic health conditions: changing prevalence in an aging population and some implications for the delivery of health care services. *Canadian Journal on Aging* 2010;29(1):11–21.
9. Inkelas M, Raghavan R, Larson K, Kuo AA, Ortega AN. Unmet mental health need and access to services for children with special health care needs and their families. *Ambulatory Pediatrics* 2007;7(6):431–8.
10. Browne G, Roberts J, Gafni A, Byrne C, Kertyzia J, Loney P. Conceptualizing and validating the human services integration measure. *International Journal of Integrated Care* [serial online] 2004 May 19; 4. [cited 21 May 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/98/195>. URN:NBN:NL:UI:10-1-100340.
11. Topping S, Calloway M. Does resource scarcity create interorganizational coordination and formal service linkages? A case study of a rural mental health system. *Advances in Health Care Management* 2000;1(1):393–419.
12. Donaldson MS. Continuity of care: a reconceptualization. *Medical Care Research and Review* 2001;58(3):255–90.
13. Strandberg-Larsen M, Krasnik A. Measurement of integrated healthcare delivery: a systematic review of methods and future research directions. *International Journal of Integrated Care* [serial online] 2009 Feb 4; 9. [cited 21 May 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/305/608>. URN:NBN:NL:UI:10-1-100547.
14. Fleury M, Mercier C. Integrated local networks as a model for organizing mental health services. *Administration and Policy in Mental Health* 2002;30(1):55–73.

15. Bernabel R, Landl F, Ganbassi G, Sgadari A, Zuccala G, Mor V, et al. Randomised trial of impact of model of integrated care and case management for older people living in the community. *British Medical Journal* 1998;316(7141):1348–51.
16. Beland F, Bergman H, Lebel P, Clarfield AM, Tousignant P, Contandriopoulos AP, et al. A system of integrated care for older persons with disabilities in Canada: results from a randomized controlled trial. *Journal of Gerontology Series A: Biological Sciences and Medical Sciences* 2006;61(4):367–73.
17. Larsen DL, Cannon W, Towner S. Longitudinal assessment of a diabetes care management system in an integrated health network. *Journal of Managed Care Pharmacy* 2003;9(6):552–8.
18. Browne G, Byrne C, Robert J, Gafni A, Whittaker S. When the bough breaks: provider-initiated comprehensive care is more effective and less expensive for sole support parents on social assistance. *Social Science and Medicine* 2001;53(12):1697–710.
19. Browne G. More effective/less expensive health services also address the social determinants of health. *Journal of the Ontario Association of Social Workers* 2011 Mar 11; 37(1). [cited 21 May 2012]. Available from: <http://newsmagazine.oasw.org/magazine.cfm?magazineid=8&articleid=143>.
20. Hebert RR, Veil AA. Monitoring the degree of implementation of an integrated delivery system. *International Journal of Integrated Care* [serial online] 2004 Sep 20; 4. [cited 20 Jul 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/106/211>. URN:NBN:NL:UI:10-1-100348.
21. Leonard BL, Parasuraman A. *Marketing services: competing through quality*. New York: Free Press; 1991. p. 20–56.
22. Wakefield DS, Halbesleben JR, Ward MM, Qiu Q, Brokel J, Crandall D. Development of a measure of clinical information systems expectations and experiences. *Medical Care* 2007;45(9):884–90.
23. Ahgren B, Axelsson R. Evaluating integrated health care: a model for measurement. *International Journal of Integrated Care* [serial online] 2005 Aug 31; 5. [cited 21 May 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/134/267>. URN:NBN:NL:UI:10-1-100376.
24. Browne G. Commentary: “Evaluating integrated health care: a system of measurement”. *International Journal of Integrated Care* [serial online] 2005 Aug 31; 5. [cited 20 Jul 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/131/262>.
25. Browne G, Kingston D, Grdisa V, Markle-Reid M. Conceptualization and measurement of integrated human service networks for evaluation. *International Journal of Integrated Care* [serial online] 2007 Dec 20; 7. [cited 20 Jul 2012]. Available from: <http://www.ijic.org/index.php/ijic/article/view/URN%3ANBN%3ANL%3AUI%3A10-1-100456/427>. URN:NBN:NL:UI:10-1-100456.
26. Brazil K, Whelan T, O'Brien MA, Sussman J, Pyette N, Bainbridge D. Towards improving the co-ordination of supportive cancer care services in the community. *Health Policy* 2004;70(1):123–31.
27. Austin PC. Using the bootstrap to improve estimation and confidence interval for regression coefficients selected using backwards variable elimination. *Statistics in Medicine* 2008;27(17):3286–300.
28. Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York: Chapman & Hall; 1993.
29. Shao J, Tu D. *The jackknife and bootstrap*. New York: Springer-Verlag; 1995.
30. Gleason JR. Algorithms for balanced bootstrap simulations. *The American Statistician* 1988;42(4):263–6.
31. Rubin DB. The Bayesian bootstrap. *The Annals of Statistics* 1981;9(1):130–4.
32. O’Gorman TW. Using adaptive weighted least squares to reduce the lengths of confidence intervals. *The Canadian Journal of Statistics* 2001;29(3):459–71.
33. Gillies RR, Shortell SM, Anderson DA, Mitchell JB, Morgan KL. Conceptualizing and measuring integration: findings from the Health Systems Integration Study. *Hospital & Health Services Administration* 1993;38(4):467–89.
34. Streiner DL, Geoff N. *Health measurement scales: a practical guide to their development and use*. 4th ed. New York: Oxford University Press; 2008.
35. Bodenheimer T. Coordinating care—A perilous journey through the health care system. *New England Journal of Medicine* 2008;358(19):1064–71.
36. Lawrence PR, Lorsch JW. Differentiation and integration in complex organizations. *Administrative Science Quarterly* 1967;12(1):1–47.
37. Sandfort JR. Exploring the structuration of technology within human service organizations. *Administration and Society* 2003;34(6):605–31.
38. Sandfort JR. Why is human services integration so difficult to achieve? *Focus* 2004;23(2):35–9.
39. Center for the Advancement of Collaborative Strategies in Health. *Partnership Self-Assessment Tool* New York: 2006 Jun. [cited 17 May 2012]. Available from: <http://partnershiptool.net>.
40. Weiss ES, Anderson RM, Lasker RD. Making the most of collaboration: exploring the relationship between synergy & partnership functioning. *Health Education and Behavior* 2002;29(6):683–98.
41. Ye C. *The Measurement of Integrated Human Services Network (The Children’s Treatment Network of Simcoe York)* Hamilton: McMaster University; 2008 Dec. [cited 17 May 2012]. Available from: <http://digitalcommons.mcmaster.ca/opensdissertations/4697>.
42. Thurston S, Paul L, Ye C, Loney P, Browne D, Browne G, et al. System integration and its influence on the quality of life of children with complex needs. *International Journal of Pediatrics* 2010 Aug 11; 2010. [cited 15 May 2012]. Available from: <http://www.hindawi.com/journals/ijped/2010/570209/>.

## APPENDIX—Statistical Methods

### Calculating the integration scores

The small letter  $n$  denotes the number of respondents. Since an agency does not rate it itself, the number of respondents is the total number of agencies less 1 when there is no non-respondent. The integration score  $\hat{x}$  is calculated by

$$\hat{x} = \frac{\sum_{i=1}^n I_i(\text{perceived involvement} = \text{expected involvement})}{n},$$

where  $I_i$  is an indicator that returns 1 if the perceived involvement is equal to the expected involvement and 0 otherwise for the  $i$ th pair of ratings. Current perception of integration can be either self- or group-perceived. Similarly, expectation can be either self- or group-expected. The variance of integration score is calculated by

$$\frac{\hat{x}(1-\hat{x})}{n},$$

where  $\hat{x}$  is the P1, P2, P3 or P4 score.

### Calculating the global integration score

#### The standard method

The capital letter  $N$  denotes the total number of agencies within a network. The average integration score estimates the global integration  $\hat{\theta}$ , given by

$$\hat{\theta} = \frac{\sum_{i=1}^N \hat{x}_i}{N},$$

where  $\hat{x}_i$  is the estimated P1, P2, P3, or P4 score of the  $i$ th agency.

The 95% confidence interval (CI) of a global integration score is calculated by

$$\left( \hat{\theta} - 1.96 \frac{s}{\sqrt{n}}, \hat{\theta} + 1.96 \frac{s}{\sqrt{n}} \right),$$

where  $s$  is the sample standard deviation of the integration scores.

#### The weighted-average method

The weighted global integration score is calculated by

$$\hat{\theta}_{\text{weighted}} = \frac{\sum_i^n \hat{x}_i w_i}{\sum_i^n w_i},$$

where  $\hat{x}_i$  is the integration score of the  $i$ th agency and  $w_i$  is the inverse of variance of  $\hat{x}_i$ .

The 95% CI of the weighted global integration score is calculated by.

$$\left( \hat{\theta} - 1.96 \sqrt{\frac{1}{\sum_i^n w_i}}, \hat{\theta} + 1.96 \sqrt{\frac{1}{\sum_i^n w_i}} \right).$$

#### The bootstrap method

In the standard bootstrap, we randomly draw an integration score from all agencies with replacement repeatedly to generate a bootstrap sample. We repeat that procedure to generate a large number of  $k$  bootstrap samples. The choice of  $k$  often depends on the computational power. In the balanced bootstrap, every integration score appears the same number of times among bootstrap samples. In other words, each original integration score is selected  $k$  times and allocated randomly in the  $k$  bootstrap samples. We achieve that by creating  $k$  replicates of original sample and then randomly sampling without replacement among those replicates to create  $k$  equal bootstrap samples. We calculate the mean of each bootstrap sample by

$$\hat{\theta}_{b_j} = \frac{\sum_{i=1}^N \hat{x}_i^b}{N}, \quad j=1, 2, \dots, k,$$

where  $\hat{\theta}_{b_j}$  is the average integration score for the  $j$ th bootstrap sample and  $\hat{x}_i^b$ 's are the integration scores selected in that bootstrap sample.

When  $k$  is large enough (usually over 1000), aggregating the means of all bootstrap samples constitutes a sampling distribution of the global score. The bias in estimating the global score can be then approximated by

$$\widehat{Bias} = \frac{\sum_{j=1}^k \hat{\theta}_{b_j}}{k} - \hat{\theta} = \hat{\theta}_b - \hat{\theta},$$

where  $\hat{\theta}_b$  is the expected global integration score from bootstrap samples and  $\hat{\theta}$  is the mean of the original sample.

The bias-adjusted bootstrap estimate of the global integration score is calculated by

$$\hat{\theta}_{\text{bias-adjusted}} = \hat{\theta} - \widehat{Bias} = 2\hat{\theta} - \hat{\theta}_b.$$

The 95% bias-corrected and accelerated (BCa) confidence interval for the global integration score is  $(m_1, m_2)$  where

$$m_1 = \Phi \left( \hat{z}_0 + \frac{\hat{z}_0 - 1.96}{1 - \hat{a}(\hat{z}_0 - 1.96)} \right) \text{ and } m_2 = \Phi \left( \hat{z}_0 + \frac{\hat{z}_0 + 1.96}{1 - \hat{a}(\hat{z}_0 + 1.96)} \right)$$

The  $\Phi(\cdot)$  is the standard normal cumulative function;  $\hat{a}$  is the acceleration; and  $\hat{z}_0$  is the bias-correction. The

method of calculating  $\hat{a}$  and  $\hat{z}_0$  empirically is provided in [28].

In the Bayesian bootstrap, each bootstrap sample generates a posterior probability for each integration score in the original sample. We draw  $N-1$  uniform random numbers between 0 and 1, order them to be  $c_1, c_2, \dots, c_N$ , and let  $c_0=0$  and  $c_1=1$ . We then calculate the gaps  $g_i=c_i-c_{i-1}$ ,  $i=1, \dots, N$  and attach  $g=(g_1, \dots, g_N)$  as the vector of probabilities to the original integration scores  $\hat{x}_1, \dots, \hat{x}_N$  for that bootstrap sample. We repeat that procedure to create  $k$  bootstrap samples. Details of the procedure can be found in [31]. The posterior estimate of the average integration scores in each bootstrap sample is calculated by,

$$\hat{\theta}_{b_j} = \sum_i^N g_i \hat{x}_i, j=1, 2, \dots, k.$$

The Bayesian bootstrap (BB) estimate of the global integration score can be calculated by the average of all posterior estimates from bootstrap samples, i.e.,

$$\hat{\theta}_{BB} = \frac{\sum_{j=1}^k \hat{\theta}_{b_j}}{k}.$$

The 95% credibility interval can be obtained from the empirical distribution of all posterior estimates  $\hat{\theta}_{b_j}$ 's.