Using the QI Maturity Tool to Classify Agencies Along a Continuum

Brenda M. Joly
University of Southern Maine, bjoly@usm.maine.edu

Maureen Booth
University of Southern Maine, maureen.booth@maine.edu

Prashant Mittal
University of Southern Maine, pmittal@usm.maine.edu

Yan Zhang
Texas Tech University, Health Sciences Center, yan.zhang@ttuhsc.edu

Follow this and additional works at: https://uknowledge.uky.edu/frontiersinphssr

🔴 Part of the Health and Medical Administration Commons, Health Policy Commons, Health Services Administration Commons, and the Health Services Research Commons

Using the QI Maturity Tool to Classify Agencies Along a Continuum

Abstract

Major investments have been made to encourage health departments to implement quality improvement (QI) efforts. Yet, there are few empirically tested tools for public health agencies that assess these efforts and classify health departments along a QI continuum. This paper presents a new classification scheme for measuring QI Maturity in public health agencies based on a validated tool. The findings can be used to establish benchmarks, make comparisons and conduct future research linking QI and population health outcomes.

Keywords

Quality improvement, health departments, measurement, classification scheme, psychometric properties

INTRODUCTION

Experts agree on the need to build the science of quality improvement (QI) in public health[1] including the development of valid, reliable and predictive measures to assess QI.[2] Yet, these measures must be practical to administer, able to inform QI efforts and capable of driving culture change. In an effort to strike a balance between these competing demands, we made two modifications to the QI Maturity Tool[2] to assure its utility while maintaining necessary psychometric characteristics. First, we shortened the tool through a second round of refinement and testing. Second, we developed an algorithm and classification scheme that health departments could use for benchmarking and comparison. The purpose of this paper is to describe our work in both of these areas and to present the five categories of "QI maturity," a theoretical concept that reflects an agency's culture, capacity and alignment of ongoing and systematic improvement efforts. This work is based on data collected from local health departments (LHDs) in 2011 as part of the Multi-State Learning Collaborative (MLC). The findings revealed five categories of QI maturity: 1) beginning, 2) emerging, 3) progressing, 4) achieving, and 5) excelling. Our findings have implications for health departments seeking to assess and monitor their QI efforts over time and for researchers who are interested in comparing overall QI maturity scores to outcomes of interest.

METHODS

The QI Maturity Tool (version 3.0) was administered to all LHDs within the 16 participating MLC states in January through March, 2011. A total of 670 agencies completed the survey, for a response rate of 57.8%. Of that total, 599 responded to 35 or more of the 37 Likert scale items in the QI tool. We repeated our initial round of psychometric testing described elsewhere[2] based on the inclusion of five additional items addressing previously reported weak or single item dimensions. The testing included an assessment of the tool's internal consistency reliability as well as its dimensionality based on Chronbach's alpha calculations and factor analysis, respectively. Questions with poor factor loadings (< 0.50) or cross loadings or those representing a single item dimension were eliminated. We also assessed the tool's convergent validity and divergent validity using the multi-trait multi-method technique. A matrix was computed with item-scale correlations which were then compared across scales. This approach helped to assess the relationship of each item with its own scale, as well as its correlations with other scales. Item convergent validity was assessed by checking the range of item-scale correlations. High item convergent validity was indicated if the item correlated noticeably with the relevant scale. A threshold of 0.40 was used in our study as suggested by Karlsson and colleagues.[3] Item divergent validity was assessed by comparing correlations between items and other scales to see if the items are stronger measures of their own construct than that of other constructs. Low item divergent validity was indicated if an item correlated significantly higher with any other scale than with its own scale. Significance was deemed if the item-scale correlation for a scale was two standard errors higher.[4] The standard error of the correlation coefficient is approximately equal to 1 divided by the square root of the sample size. In our study, two standard errors is equal to: $2(1/\sqrt{599}) = 0.082$.

Average scores were calculated for each domain and the underlying dimensions. These scores were then assessed in an algorithm for categorizing the LHDs. Due to the inherent variation within domains and an imperfect one on one correspondence between QI domains and practice, it was not possible to create a simple summative algorithm that would fit every LHD perfectly. According to the strong linear positive association between higher domain composite scores and higher levels of practice among LHDs observed in our cluster analysis, we created a hierarchically incremental algorithm to generate five non-overlapping categories taking into consideration two

factors: 1) the use of a probabilistic deciles method to capture the essence of the majority of LHDs in that category, and 2) a hierarchical method imposing stricter criteria as an LHD moved towards higher categories. The methodology is very similar to multivariate regression modeling where the best fitted line defines and predicts the center of the data.

RESULTS

The results of the psychometric testing are provided in Table 1. Chronbach's alpha estimates ranged from 0.75 to 0.87 suggesting moderate to good internal consistency reliability across all domains. Dimensionality testing resulted in the elimination of two items from the QI organizational culture domain, one item from the QI capacity and competency domain, four items from the alignment and spread domain and one single-item dimension. Additionally, the results revealed that item-internal consistency (convergent validity) is satisfactory and the inclusive criterion of a correlation of 0.40 or higher was met for all items. Finally, the divergent validity test among all three factors was satisfactory since none of the items exhibited higher correlations with other scales in comparison with the proposed scale (by two standard errors or more).

As seen in Table 2, five categories of QI Maturity ranging from immature to highly mature include: 1) beginning, 2) emerging, 3) progressing, 4) achieving, and 5) excelling. The table shows the upper limit of the first category (beginning) was chosen to be the 60th percentile for that category indicating 60% of the "beginning" LHDs scored 99 or less. Similarly, 70% of the LHDs scored less than 106 in the second category and so on. As the composite of practice and domain summary scores got higher, the criteria for placing an LHD got stricter as well (e.g., an agency needed to score disproportionately higher in order to be included in the next category). The cumulative membership of the five categories was very close to the actual cumulative membership in the data. Approximately 4% of respondents were classified as "excelling", representing agencies achieving high levels of QI sophistication and a pervasive culture of QI. Ten percent of agencies were classified as "achieving" which suggests fairly high levels of QI practice, a commitment to QI, and an eagerness to engage in the type of transformational change described by QI experts.[5] Over one-third (36%) of LHDs were identified as "progressing" agencies. These LHDs typically have some QI experience and capacity but often lack commitment, have minimal opportunities for QI integration throughout the agency and are less sophisticated in their application and approach. Of the 18% of agencies classified as "emerging," there is evidence to suggest that these LHDs have newly adopted QI approaches, albeit with limited capacity. Emerging agencies have limited QI culture and few, if any, examples of attempts to incorporate QI as a routine part of practice. Finally, "beginning" LHDs are those that have typically not yet adopted formal QI projects, applied QI methods in a systematic way, or engaged in efforts to build a culture of QI. The percentages of agencies in each category described above mirrored the actual QI practice status based on our observations with the MLC national evaluation designed, in part, to assess QI efforts.

IMPLICATIONS

The QI Maturity Tool is the only known validated tool to assess and monitor QI efforts in public health agencies and the streamlined tool and accompanying classifications have several implications for practitioners and researchers. First, the tool provides agencies with a practical and sound approach for measuring various aspects of QI maturity. Given our new classification scheme, agencies (based on a single respondent) are now able to relatively easily: 1) calculate their score by summing the 29 items and 2) categorize their current level of QI maturity based on the cut-offs provided in Table 2. For instance, if an agency scores a total of 95, then it would be ranked as

"beginning." Second, the accompanying scores and classification scheme allow agencies to establish benchmarks and monitor their level of organizational QI maturity over time. Third, the tool has shown promise for measuring increased organizational QI capacity and it may therefore be of interest to agencies that are seeking to assess and document their QI investments in training, technical assistance, and project implementation, particularly if researchers are able to develop domain-level sub-scores.

Finally, if researchers are able to develop a standard process for administering this tool as well as a central repository for this data, our field would be better positioned to: 1) determine if differences in QI scores exist among certain groups of interest (e.g., accredited, non-accredited agencies), 2) assess the tool's predictive capability, and 3) explore the role of QI Maturity on select outcomes of interest. Important next steps include the validation of the new categories and identification of the most appropriate administration process (e.g., single versus multiple respondents) for the QI Maturity Tool.

SUMMARY BOX

What is already known on this topic?
Over the last several years, significant investments have been made to promote the adoption and acceleration of quality improvement (QI) efforts in health departments. Given these investments and the growing emphasis in this area, there is a need to create tools that measure changes based on various aspect of QI.

What is added by this report?
While existing public health efforts and literature explore characteristics and opportunities for building a QI culture, this study provides a classification scheme for measuring QI Maturity based on empirical work with a validated tool measuring multiple domains of interest.

What are the implications for public health practice/policy/research?
Implications for practice – this tool and accompanying classification scheme can be used for self-assessment purposes and as an accountability mechanism with leaders and policy makers.
Implications for research – this study provides important evidence that different levels of QI maturity exist among health departments in the United States. The five levels of QI maturity have been empirically developed, yet they now need to be validated. Additionally, more efforts are needed to understand how progression along a QI continuum influences public health

---

REFERENCES

[1]Riley W, Lownik B, Halverson P, Parrotta C, Godsall JR, Gyllstrom E, Gearin KJ, Mays G. (2012). Developing a taxonomy for the science of improvement in public health. J Public Health Manag Pract 2012;18 (6): 506-14.

[2]Joly BM, Booth M, Mittal P, Shaler G. Measuring Quality Improvement in Public Health: The Development and Psychometric Testing of a QI Maturity Tool. Eval Health Prof 2012; 35(2): 119-147.

[3]Karlsson J, Persson L-O, Sjöström L, Sullivan M: Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. *Intern J Obesity* 2000, 24:1715-1725.

[4]Ware JE, Gandek B: Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *J Clinical Epi* 1998, 51:903-912.

[5]Riley WJ, Parsons HM, Duffy GL, Moran JW, Henry B. (2010). Realizing transformational change through quality improvement in public health. J Public Health Manag Pract 2010; 16(1): 72-8.

Table 1. QI Maturity Tool Factor Loadings, Chronbach's Alpha Estimates and Alignment with Hypothesized Domains (n= 599)

Domains and Items	Factor Loading	Dimension	Chronbach's Alpha	Convergent Validity # of Inter-Item Correlations $\geq$ 0.40	Divergent Validity # items that have higher correlation with other scale <0.082
QI Organizational Culture					
Leaders receptive to ideas for improving quality	.712	Commitment			
Impetus for improving quality is internal	.694	Commitment	0.75	3 of 3	0 of 3
Leaders work together for common goals	.786	Commitment			
~~Agency data shared for improvement purposes~~	~~.228~~	~~Collaboration~~			
~~Matters reviewed in respectful way~~	~~.449~~	~~Collaboration~~			
Staff help solve problems	.934	Collaboration	0.80	2 of 2	0 of 2
Staff routinely contribute to decisions	.756	Collaboration			
QI Capacity & Competency					
Leaders are trained in basic QI methods	.724	Skills	0.85	2 of 2	0 of 2
Staff members are trained in basic QI methods	.941	Skills			
Staff has skills to monitor quality of programs	.589	Methods			
Agency has objective quality measures	.769	Methods			
Staff uses methods to identify root causes	.508	Methods	0.87	6 of 6	0 of 6
Staff uses best or promising practices	.643	Methods			
Programs are continuously evaluated	.932	Methods			
The quality of programs is routinely monitored	.829	Methods			
~~Established process exists for identifying QI priorities~~	~~.484~~	~~Investment~~			
Agency has a Quality Improvement Officer	.747	Investment			
Agency has QI Council, Committee or Team	.903	Investment	0.85	3 of 3	0 of 3
Agency has QI Plan	.791	Investment			
QI Alignment & Spread					
Job descriptions include QI responsibilities	.799	Integration			
Staff is aware of external QI expertise	.629	Integration			
Staff at all levels participate in QI	.651	Integration			
Customer satisfaction information routinely used	.699	Integration	0.86	7 of 7	0 of 7
QI efforts usually adopted by other programs	.657	Integration			
Accurate and timely data available for QI	.646	Integration			
Improving quality is integrated into agency practice	.705	Integration			
~~Agency allocates sufficient time for QI~~	~~.505~~	~~Integration~~			
Staff has authority to make change	.688	Authority	0.78	2 of 2	0 of 2
Staff has authority to work across program boundaries	.921	Authority			
~~Managers are accountable for improvement~~	~~.124~~	~~Implementation~~			
~~Implementing QI is challenging among staff~~	~~.088~~	~~Implementation~~			
~~QI approaches are compatible with activities~~	~~.201~~	~~Implementation~~			
Spending time and resources on QI is worth it	.834	Value			
Key decisions makers think QI is important	.678	Value	0.85	4 of 4	0 of 4
Using QI will impact health of my community	.897	Value			
Staff will notice change as a result of QI	.696	Value			
~~Staff have adequate time and support for QI~~	~~.853~~	~~Single item dimension~~			

Note: Strikethrough represents items that were deleted due to a low factor score (<0.50), cross-loading with one or more dimensions or due to a remaining single item dimension.

Table 2. QI Maturity Scores by Category (n= 599)

QI Maturity Categories	Score Range (29-145)	Percent of Respondents	Upper Limit Percentile of the Group
Beginning	≤ 99	32.0	60th
Emerging	100-106	17.8	70th
Progressing	107-120	35.7	80th
Achieving	121-139	10.4	90th
Excelling	≥ 140	4.0	100th