

| Article details: 2017-0166 | |
|--|---|
| Title | Accelerating surgical quality improvement in Ontario through a regional collaborative: a quality improvement study |
| Authors | Timothy Jackson MD MPH, David Schramm MD MSc, Husein Moloo MD MSc, Lee Fairclough MRT MHSc, Azusa Maeda PhD, Tricia Beath PT MHA, Avery Nathens MD PhD |
| Reviewer 1 | Dr. Erik Skarsgard |
| Institution | Department of Surgery, BC Children's Hospital, Vancouver, BC |
| General comments (author response in bold) | <p>One concern relates to the validity of measurements that are used to justify statements like "Our initial findings suggest that the support provided by ON-SQIN improved capacity for QI and ability to respond to their own data".</p> <p>1. Do you have any evidence beyond what is provided in Table 2 (self-reporting on vague capacity measures, such as "organizational engagement in surgical QI") that actually confirms this? Objective measures might include evidence of implementation of an ERAS program, other QI bundles, a CUSP program, evidence of surgeon completion of on-line educational programs, institution of safety rounds, creation of pre-printed order sets, monthly reporting of occurrences etc. To enhance quality improvement capacity, members of the collaborative were provided access to various resources as highlighted in the methods section (lines 102-107: "Hospitals were provided with access to resources and tools including site visits, an established quality improvement program...."). We revised the manuscript to include the number of sites that participated in such programs in the Results section as an objective measure of increased capacity (lines 161-164: "Resources provided by the Surgical Network were utilized over the 18-month period, including the online platform (>200 visits per month), Improving and Driving Excellence Across Sectors (7/18 sites), Open School e-Learning module (15/18 sites), enhanced recovery after surgery programs (15/18 sites), and Canadian Surgical Site Infection Prevention Audit (8/18 sites)".).</p> <p>2. It is also unclear what additional human resources for quality improvement activities (other than one or two SCRs that come with the NSQIP program) may exist across the various hospitals in the collaborative to increase their QI capacity. Some of the UHN hospitals may have personnel dedicated to improvement charters, process redesign and change management, while smaller hospitals likely have little or no dedicated resources for this. Does this inequity exist and how does the collaborative address this? The collaborative structure provided support through the Surgical Network Steering Committee, the Network Delivery Team, and the Provincial Surgical Lead to support quality improvement activities. In addition, Health Quality Ontario (HQO) provided financial support to help offset the initial costs for the first 18 months of the program, tailored to the needs of individual hospital assessed by the initial questionnaire. We revised the manuscript to discuss this (lines 107-110: "In addition, hospitals were provided with financial support to help offset the initial costs for the first 18 months of the program....", lines 239-242: "Hospitals were provided with financial support during the first 18 months").</p> <p>3. The timing of unadjusted outcomes analysis, and the assumption that improvements seen could be attributed to the Collaborative in hospitals that had never received an SAR should be questioned. Until a new NSQIP program participant receives that first SAR, its program leaders are "in the dark" as to how they are performing, and it is usually that first benchmarked report, that provides the engagement opportunity for the front line providers and administration to support QI measures. I think this should be addressed in the discussion (see also #7 below). On the other hand, the 5 hospitals that were already part of NSQIP (and therefore further along on their "QI journey") were disproportionately representative of large academic hospitals. Hospitals collected NSQIP data at least for 4 months prior to determining their targets. While the first SAR is the gold standard to determine areas that need improvement, the on-demand data is readily available and can be used at an early stage to identify opportunities for improvement. We revised the manuscript to address this (ref #24, lines 227-237: "Hospitals collected data at least for 4 months prior to determining their targets for improvement....."). As pointed out, the aggregate data may be overrepresented by large academic hospitals. However, as discussed in the limitations section, we were unable to consider the size of the hospital since the collaborative received de-identified aggregate data at the time of the analysis. We revised the manuscript to discuss this limitation in the current data analysis and our current approach to overcome this limitation (i.e. access to unblinded data) (lines 265-272, "Limitations of the use of aggregated data include.....).</p> <p>4. What was the relative contribution of aggregate cases from existing versus new NSQIP program participants? As the collaborative received blinded, aggregated data, we were unable to determine the relative contribution of existing v.s. new participants. We discuss in the limitation section that we were unable to consider the impact of hospital size and different degrees of existing quality improvement capacities. As of January 2018, the collaborative is unblinded, such that we can examine data at an individual site level. We revised the manuscript to discuss this limitation in the current data analysis (lines 265-272, "Limitations of the use of aggregated data include.....).</p> <p>5. How was the Collaborative able to drive change improvements in hospitals that had no idea of whether they were a low or high performer for surgical site or urinary tract outcomes? Would the results be different if the aggregate occurrence analysis was conducted for new (n=13) versus existing (n=5) NSQIP programs? Hospitals collected their NSQIP data at least for 4 months prior to determining their targets, which were reported in their SQIP (lines 227-228: "Hospitals collected data at least for 4 months prior to determining their targets for improvement"). By comparing their baseline data to that of the collaborative and/or NSQIP hospitals (raw or on-demand data), each hospital was able to identify targets for improvement. Hospitals were then able to use resources specific to those targets that were made available by the Surgical Network. While the results may have been different if the aggregate analysis was conducted for new v.s. existing hospitals, we were unable to perform the analysis as we only received de-identified, aggregate data at the time. We revised the manuscript to address this limitation (lines 265-272, "Limitations of the use of aggregated data include.....).</p> <p>6. On page 9, line 193, it states "...hospitals reported either their raw data or their risk-adjusted data". I'm unclear on how hospitals could report their risk-adjusted data, unless you are referring to the Odds ratio (observed/expected) for a given occurrence, even though that was not part of the current analysis. This could be clarified. Few hospitals reported their Odds ratio based on the SAR. However, as pointed out by the reviewer, this should not be part of the current analysis. We revised the manuscript to exclude those hospitals that reported Odds ratio and only included those that reported raw data for clarification (Figure 1, lines 144-146: "Hospitals that reported OR in their Surgical Quality Improvement Plan and those that did not submit the year-end Surgical Quality Improvement Plan were excluded from the analysis."). To further clarify, we revised the figure to present the self-reported rates of SSI/ UTI for each hospital.</p> <p>7. Figure 3 shows changes in self-reported rates of SSI, UTI in hospitals who had self-identified as targeting either SSI reduction, UTI reduction</p> |

| | |
|--|---|
| | <p>or both. To be clear, these are just average hospital rates for 3 successive 6 month periods? Can these improvements be specifically correlated with process changes at these hospitals? A sample control chart showing UTI occurrences and the timing of interventions might be a useful figure to include.</p> <p>As discussed above, we revised the figure to present the self-reported rates of SSI/ UTI for each hospital for clarification. While the SQIP indicates the information on change ideas implemented at each hospital, it is difficult to demonstrate the change in postoperative occurrences with respect to the change ideas for all of the hospitals. To address the comment, we included a reference in the revised manuscript, demonstrating change in UTI occurrence following an implementation of a QI bundle at one of the hospitals in the collaborative (ref #22, lines 194-197: "Many hospitals identified reduction of surgical site and urinary tract infections as their targets for improvement and implemented best practice bundles (22), and demonstrated reduction of these postoperative occurrence".).</p> <p>8. While I realize that the focus of the paper is on the Collaborative, which uses unadjusted data for performance measurement, the discussion (interpretation) does not adequately address the importance of appropriately risk-adjusted outcomes comparisons. The shading of the bars in Figure 2 is difficult to make out, but it looks like more than 70% of the ON-SQIN hospitals "need improvement" in SSI performance. Some comment should be made that describes the relative benefits of adjusted versus unadjusted data in surgical QI activities.</p> <p>We thank the reviewer for his suggestion. The raw data is informative for frontline staff at the beginning of the quality improvement initiative as it can be easily interpreted and gives the perspective of the number of patients affected. Hospitals can also obtain the on-demand, risk-adjusted rates report. While it is not as informative as the risk-adjusted data in the SAR, the on-demand data is readily available and can be used at an early stage to identify opportunities for improvement as the data comparison period can be specified. Nevertheless, SAR is still the gold standard with respect to accuracy (ref #24). We revised the manuscript with appropriate references to discuss the benefits of adjusted and unadjusted data (lines 227-237: "Hospitals collected data at least for 4 months prior to determining their targets for improvement.....").</p> <p>9. Page 10, line 229 reads: "Few rural and small ON-SQIN hospitals were unable to maintain their participation in the ACS NSQIP due to the associate(d) cost...". Does this mean there has been hospital drop out from the Ontario NSQIP program? Three sites (community and small/ rural hospitals) were not able to continue with NSQIP beyond the funded 18-month period. However, continued participation in the collaborative activities outside of NSQIP, including conferences and access to resources, remains viable for those hospitals. We revised the manuscript to discuss this (lines 245-257: "However, 3 rural and small hospitals were unable to maintain their participation in the National Surgical Quality Improvement Program due to the associate cost, highlighting challenges to sustain and disseminate the collaborative.").</p> <p>10. Some mention of measures taken to ensure privacy protection within the collaborative should be provided. The NSQIP participation agreement ensures that the data reported back from NSQIP does not identify individual hospitals. Similarly, in the personalized SAR for the Surgical Network, hospitals were only able to identify their results among the full distribution of non-identified Surgical Network hospitals, allowing for benchmarking within the group. This agreement was recently revised such that the participants can benchmark and identify other hospitals, especially the "high-performers", to encourage sharing of successful quality improvement initiatives. We revised the manuscript to discuss this (lines 268-272: "To address these limitations, the Surgical Network has recently revised the data sharing agreement to collect unblinded hospital-level data...").</p> |
| Reviewer 2 | Dr. Steven Lopushinsky |
| Institution | Toronto, Ont. |
| General comments (author response in bold) | <p>1. The introduction of the paper describes NSQUIP and references regional collaborations in both the United States and Canada. Specifically there is reference to a similar collaboration in British Columbia but no further discussion about its outcomes and lessons learned. It is not clear then what the paper adds to the literature or how it compares to those programs described before them. Perhaps the discussion could be reframed to reflect process of developing a regional collaborative (i.e. lessons learned) particularly in the context that the quality improvements demonstrated are relatively small; for example, while UTI relative risk was statistically significant at 0.77, the absolute risk reduction is about 0.6% and there is little mention to how clinical meaningful this is.</p> <p>We thank the reviewer for his suggestion. We revised the manuscript to focus more on the process of developing a regional collaborative, as well as aspects that are unique to the Ontario collaborative (Methods section: "Collaborative structure", "Quality improvement resources and community of practice", lines 140-142: "Lastly, self-reported outcomes were collected from member hospitals through a Surgical Quality Improvement Plan...", 130-132, 207-223: "Although it is difficult to identify specific components of the collaborative that contribute to improvement, we believe that the collaborative structure, including the community of practice, provides a model for spread, scale, and sustainability across member hospitals.....").</p> <p>2. The term "community of practice" is introduced and is perhaps not described in a meaningful enough way to non-quality improvement researchers. There is precedent to its use for surgical outcomes in Ontario, specifically in surgical oncology through Cancer Care Ontario (Fung-KeeFung et al. Q Manage Health Care 2008;17(2):174-185) that could be used as historical context for the current model of implementation.</p> <p>We thank the reviewer for his suggestion. We included references and discuss the definition and role of community of practice in more detail (ref #15 and 23, lines 78-80: "...support surgical teams to accelerate improvement through a community of practice, defined as...", lines 210-213: "The community of practice was built upon a previously described model...").</p> <p>3. A fundamental issue requiring discussion is the impact of the regional collaborative on improved surgical outcomes versus those that could be applied at the local hospital level through bundles of care or other directive interventions. Although a future cost-benefit analysis is mentioned, it is not clear from this paper how conference calls, webinars, and in-person meetings to develop this "community of practice" contributed to meaningful improved surgical quality. The literature certainly demonstrates multiple examples of rapid improvement in indicators, such as SSI and UTI, with NSQUIP data availability and/or other hospital-driven programs (i.e. bundles of care, ERAS) without the enhanced community piece.</p> <p>As pointed out by the reviewer, it is difficult to discriminate the impact of the Surgical Network versus NSQIP or other hospital-driven quality improvement initiatives alone. We refer to studies in the Introduction and Discussion sections to highlight that while participation in NSQIP has been clearly associated with a reduction of postoperative adverse events, participation in a collaborative may further accelerate improvements (ref#7,12). We believe that the collaborative structure provides a model for spread, scale, and sustainability across participating hospitals. We revised the manuscript to address this comment (lines 71-74: "Notably, a study demonstrated that...", lines 205-223 "While participation in the National Surgical Quality Improvement Program has been clearly associated with....").</p> <p>4. The statistical analysis section should be further developed as the only two tests described are calculating percent change and relative risk, neither of which necessarily have an associated pvalue. Further, a number of times through the paper the authors cite "a trend" for</p> |

improvement although confidence intervals overlap suggesting no difference. I would argue that a true test for trend (i.e. through a regression analysis or other) be performed if such statements are to be made.

The data on relative risk and % change (Table 4) has an associated p-value. The statistical method used here is equivalent to that of previous reports on similar outcomes (ref #10). Figure 1 (previously Figure 3) was revised to present the self-reported rates of SSI/ UTI for each hospital, and the statement regarding "trend" for improvement was removed.

Minor

1. The abstract only focuses on the positive findings (ARF and UTI) and neglects the negative findings of SSI

Reductions of ARF and UTI were the only variables that were statistically significant. Thus, we did not include the findings on SSI in the abstract.

2. In terms of manuscript structure, I would suggest that "Building the collaborative" and "Establishing the Community of Practice" are part of the methods

We thank the reviewer for his suggestion. "Building the Collaborative" section was moved to the Methods and Interpretation section, and Table 1 was removed as suggested by the editors. Some parts of the "Establishing the Community of Practice" section was moved to the Methods section.

3. Reference to the number of participating hospitals (page 7 lines 141-143) expanding to 33 seems extraneous as not part of the studied population (n=18, restricted to those starting in 2015) -- reasonable to include as part of the discussion however

We revised the manuscript to move this reference to the Interpretation section (lines 200-204: "Since 2015, the number of member hospitals grew to 33 by the end of 2016, and 11 of the 14 Ontario Local Health Integration Networks have hospitals participating in the Surgical Network.").

4. With respect to "Building Capacity" (page 8, lines 149-157), who at each centre completed the survey? The response may be biased if a single person is completing the survey (i.e. surgeon champion), and came on in 2015 for the start of the project and was actively involved through 18-months, thus having an exposure bias (may not be aware of other activities at their hospital preceding their involvement in this project). If indeed a single response per institution, is it representative of all views at that institution.

The assessment was completed by Surgeon Champions, Surgical Clinical Reviewers, quality improvement team members and/or surgical administrative staff at each hospital. Thus, we believe that the response represents the view of the key members involved in quality improvement at each institution. We revised the manuscript to add details of the questionnaire in the Methods and Results section.

5. Further with respect to the survey, would the categories of "Active surgical QI initiative" and "Organizational engagement in surgical QI" be self-fulfilling if each of the centers took part in this collaboration? In this survey, how as "Collaboration within surgical team" defined? Also, would this be impacted by other initiatives like surgical safety checklist audits that may be happening simultaneously as part of other quality improvement efforts?

The details of the survey was added to clarify the definitions (Appendix 1). The answers may certainly be impacted by other initiatives such as audits performed as part of the Surgical Network or other hospital-led initiatives. We revised the manuscript to discuss this (lines 220-223: "Although other hospital-led quality initiatives may exist independent of the collaborative....").

6. In the Results section "Impact on Surgical Outcomes," for the first time the mention of acute renal failure is identified and measured. Was this decided a priori or were other morbidities considered? Interesting that acute renal failure was demonstrated to have a relative risk of 0.48, not fully explained by the drop in UTI relative risk (RR 0.77). This may point to flaws of using non-adjusted data (i.e. account for co-morbidity) and also that of random chance in findings. Similarly, is there a potential explanation for increased incidence of organ/space SSI, wound disruption, and progressive renal insufficiency as there is no further comment on this in the discussion.

In addition to UTI and SSI, other postoperative occurrences (e.g. respiratory, CNS, cardiac, and others) were captured as part of the ACS NSQIP. However, we only included urinary and wound occurrences as there was a room for improvement in these indicators (lines 171-173: "The performance of the Surgical Network hospitals on postoperative indicators demonstrated that..."). The limitation of the study is that the reason for change is not readily obvious, as discussed elsewhere (ref #10). We revised the limitations section to address this (lines 267-272: "Furthermore, reasons for aggregate change in postoperative occurrences is not readily obvious..."). We anticipate that we would be able to identify successful initiatives that led to improvements with unblinded data. We did not discuss the observations of the increased incidence of organ/space SSI, wound disruption, and progressive renal insufficiency as they were not statistically significant.

7. It is a flaw to report improvements in overall occurrence of superficial incisional SSRI and deep incisional SSI when the confidence interval crosses one (page 9, lines 188-190). To this end, I would argue statements in the interpretation are overly strong / conclusive that there is a demonstrated improvement.

We removed the sentence and revised the manuscript to tone down on the interpretation and focus more on the process of building the collaborative as suggested by the reviewer.

8. Figure 2: Incidence of postoperative occurrences in "general, vascular, colorectal, all surgeries..." Can you please clarify as to whether "all surgeries" were captured or just the previously mentioned specialties?

Not all hospitals capture data on all surgeries performed at their hospitals. We revised the manuscript to clarify this (lines 131-132: "13 indicators for general, vascular, colorectal, or all cases surgeries between January and December 2015").

9. Regarding performance on postoperative indicators, can the authors provide a pre-and post-performance comparison (i.e. is there a change in percentage of "needs improvement" or "exemplary" outliers). Changes may suggest a broader change in care, culture of quality improvement and/or patient population (unadjusted data) which may or may not directly reflect the targeted QI efforts in this study.

We are unable to compare the performance ("needs improvement" status etc.) for the initial 18 hospitals beyond December 2015 as additional hospitals joined the Surgical Network in 2016 and the data is aggregated. To overcome this limitation, the Surgical Network has now shifted to obtaining unblinded data, which can be used to determine the level of changes at each hospital. We discuss this in the limitations section (lines 268-272: "To address these limitations, the Surgical Network has recently revised the data sharing agreement to collect unblinded hospital-level data...").