A hospital-randomized controlled trial of a formal quality improvement educational program in rural and small community Texas hospitals: one year results

GIOVANNI FILARDO1, DAVID NICEWANDER1, JEPH HERRIN2, JANINE EDWARDS3, PERCY GALIMBERTTI3, MARI TIEZTE4, SUSAN MCBRIDE4, JULIE GUNDERSON1, ASHLEY COLLINSWORTH1, ZIAD HAYDAR1, JOSIE WILLIAMS3 AND DAVID J. BALLARD1

1Institute for Health Care Research and Improvement, Baylor Health Care System, 8080 N. Central Expressway, Suite 500, Dallas, TX 75206, USA, 2Section of Cardiovascular Medicine, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA, 3Texas A&M University Health Science Center, 3833 Texas Avenue, Suite 150, Bryan, TX 77802, USA, and 4Dallas Fort Worth Hospital Council, 250 Decker Drive, Irving, TX 75062, USA

Abstract

Objective. To investigate the effectiveness of a quality improvement educational program in rural hospitals.

Design. Hospital-randomized controlled trial.

Setting/Participants. A total of 47 rural and small community hospitals in Texas that had previously received a web-based benchmarking and case-review tool.

Intervention. The 47 hospitals were randomized either to receive formal quality improvement educational program or to a control group. The educational program consisted of two 2-day didactic sessions on continuous quality improvement techniques, followed by the design, implementation and reporting of a local quality improvement project, with monthly coaching conference calls and annual follow-up conclaves.

Main Outcome Measures. Performance on core measures for community-acquired pneumonia and congestive heart failure were compared between study groups to evaluate the impact of the educational program.

Results. No significant differences were observed between the study groups on any measures. Of the 23 hospitals in the intervention group, only 16 completed the didactic program and 6 the full training program. Similar results were obtained when these groups were compared with the control group.

Conclusions. While the observed results suggest no incremental benefit of the quality improvement educational program following implementation of a web-based benchmarking and case-review tool in rural hospitals, given the small number of hospitals that completed the program, it is not conclusive that such programs are ineffective. Further research incorporating supporting infrastructure, such as physician champions, financial incentives and greater involvement of senior leadership, is needed to assess the value of quality improvement educational programs in rural hospitals.

Keywords: quality improvement/quality management, quality indicators/measurement of quality, rural or urban/specific populations, hospital care, setting of care, training/education, human resources

Address reprint requests to: David J. Ballard, Institute for Health Care Research and Improvement, 8080 N. Central Expressway, Suite 500, Dallas, TX 75206, USA. Tel: 214-265-3670; Fax: 214-265-3640; E-mail: dj.ballard@baylorhealth.edu
What this paper adds

(1) What is already known?
- Continuous quality improvement has become a popular strategy for improving performance on health care quality indicators.
- Educational programs teaching continuous quality improvement skills and strategies have shown some success in improving quality performance in large hospitals and health care systems, as well as ambulatory care physician practices, but have not been evaluated in rural health care settings.

(2) What this study adds?
- This hospital-randomized controlled trial investigated the effect of an educational program teaching continuous quality improvement skills and strategies on performance on quality indicators for pneumonia and heart failure care in rural and small community hospitals in Texas.
- Our results showed no greater improvement in performance with the continuous quality improvement training.
- Continuous quality improvement training in the absence of organizational support such as involvement of leadership in quality improvement, physician champions to lead and promote improvement efforts and financial incentives linked to performance on quality indicators may be insufficient to effect meaningful improvements in quality of care.

Methods

The design and implementation of this trial have been detailed previously [18, 19]. Briefly, the study consisted of two phases. The first phase (September 2004 to July 2005) implemented a web-based quality benchmarking and case-review tool (Cognos PowerPlay [20]) at 64 rural and small community Texas hospitals. This tool facilitates the analysis of quality and safety measures by accessing a database of Agency for Healthcare Research and Quality indicators for hospital-to-hospital comparisons and internal case-level analysis. Each site's technology needs were assessed and deployment of the tool tailored accordingly. The hospital's reporting of required quality measures was simultaneously assessed, both with respect to which measures were reported and the methods used to collect and submit the data. In March 2006, the enrolled hospitals meeting the eligibility criteria for the second phase (submission of pre-trial core measures data and signed consent for participation from the hospital Chief Executive Officer/President [18]) were randomized either to receive the quality improvement educational program or not. The randomization scheme accounted for hospital volume of eligible congestive heart failure and community-acquired pneumonia cases and for the hospital compliance rate for the five related process measures. The use of the web-based quality benchmarking and case-review tool continued to the end of the study (September 2007) for both groups.
Quality improvement educational intervention

At the beginning of the program, participants were asked to identify the most urgent quality improvement and monitoring needs for their hospital, so that they could apply the general strategies taught to these specific problems. Content and structure of the program are described elsewhere [18]. The course consisted of two 2-day sessions teaching continuous quality improvement techniques, followed by 3 months during which participants conducted quality improvement projects, while receiving coaching via monthly teleconferences and e-mail. Participants were encouraged to focus their projects around congestive heart failure and/or pneumonia care. Projects were evaluated by course instructors. Annual conclaves, during which participants reported and discussed progress and additional initiatives undertaken, were held over the following 2 years.

Outcome measures

Primary outcome measures were Centers for Medicare and Medicaid Services core measures for congestive heart failure: left ventricular function assessment, angiotensin-converting enzyme inhibitor or angiotensin II receptor blocker for left ventricular systolic dysfunction and pneumonia: oxygenation assessment, pneumococcal vaccination, antibiotics within 4 h. We examined both individual and composite measures for each condition. In the composite measures, the denominator represents all patients eligible for at least one of the indicators, whereas the numerator represents the patients who received all the condition-specific indicators for which they were eligible. We also examined an overall composite measure for congestive heart failure, pneumonia and acute myocardial infarction (using the five acute myocardial infarction core measures: aspirin at arrival; aspirin at discharge; angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction; beta blockers at arrival and beta blockers at discharge). We did not examine acute myocardial infarction measures separately as few hospitals treated sufficient numbers of patients.

Minimum detectable differences

Using methods appropriate for cluster-randomized studies [21], minimum detectable differences were estimated for each outcome using alpha and power at 0.05 and 0.80, respectively. The absolute estimates ranged from 4.3% (oxygenation assessment) to 16.2% (pneumococcal vaccination) [18].

Data collection

Centers for Medicare and Medicaid Services core measures data were collected from Texas Medical Foundation via the state’s QualityNet Exchange system. Retrospective data (starting 1 year prior to implementation) on quality indicators were collected for the baseline assessment. Follow-up data were collected for 1 year following completion of the educational program.

Statistical analysis

The effectiveness of the educational intervention was assessed by comparing the change over time in the selected Centers for Medicare and Medicaid Services measures between study groups. Analysts were blinded to group allocation. For control group hospitals, the baseline period was the quarter of randomization (first quarter of 2006) plus the three previous quarters. The baseline period for hospitals randomized to receive the educational intervention was the quarter during which the intervention began plus the three previous quarters. Follow-up consisted of the four quarters immediately following the respective baseline periods.

To compare the change over time between study groups, we estimated a generalized linear mixed model with logit link and binomial error distribution for each measure:

$$\text{Logit}(P) = \mu + \beta_1 \text{Group} + \beta_2 \text{Time} + \beta_3 \text{Group} \times \text{Time} + \gamma_1 \text{Hospital} + \gamma_2 \text{Hospital} \times \text{Time}$$

Here, the logit of the probability of receiving the indicated treatment—or fulfilling all criteria for the composite measure for pneumonia or congestive heart failure—($P$) is expressed as a function of group membership (Group), time (baseline; follow-up) (Time), the group by time interaction (Group $\times$ Time), a random hospital effect (Hospital) and a random hospital by time effect (Hospital $\times$ Time). Because these data were from a randomized trial, we considered no further covariates for the Phase 2 models. The random hospital and hospital by time effects were included to account for correlation in the data.

Intervention effectiveness was assessed by testing the group by time interaction with an F-statistic. All models were fit using PROC GLIMMIX in SAS software (SAS Institute Inc., Cary, NC).

Results

Of the 64 hospitals enrolled for Phase I, 47 met the criteria for randomization for Phase II, with 24 being assigned to the control group and 23 to the educational intervention group. Of these, 2 hospitals did not report baseline or follow-up data, leaving 45 hospitals in the final analysis. In the educational intervention group, representatives from 16 hospitals completed the core educational program (three sessions and a quality improvement project) and representatives from six participated in the full program (educational sessions, quality improvement project, monthly coaching sessions and annual conclave) (Fig. 1). As described elsewhere, few participating hospitals sent the physician leaders, nursing leaders, administrative leaders...
and patient safety officers for whom the educational program was intended [19].

Table 1 shows the median volume of admissions eligible for each study measure during the 12-month baseline and follow-up periods for each study group. Table 2 shows the results of the intention-to-treat effectiveness analysis. Descriptive statistics and the P-values related to the intervention-group-by-time $F$-tests for the congestive heart failure measures, pneumonia measures and the ‘all-condition’ composite are shown. Improvement from baseline to follow-up was seen in both groups for all pneumonia measures and the all measure composite. For the congestive heart failure measures, the control group showed improvement whereas the intervention group showed slightly decreased performance. No significant differences in improvement were seen at 1 year follow-up between the study groups. Similar results were seen when only the 16 hospitals that completed the didactic portion of the

Figure 1 Flow diagram showing the randomization of hospitals to study arms in the randomized controlled trial testing a quality improvement educational program in rural and small community hospitals in Texas (2005–2007), as well as the level of exposure to the intervention among hospitals randomized to the intervention arm. Asterisk denotes completed the educational sessions, participated in monthly coaching conference calls and attended the 12-month follow-up conclave.
Discussion

We examined the impact of a quality improvement educational program in rural Texas hospitals in a randomized, controlled trial following implementation of a web-based quality benchmarking and case-review tool. At 12 months, no significant differences in improvement on heart failure, pneumonia, or composite quality indicators were observed between the intervention and control groups. The 47 participating hospitals represent 25% of all Texas hospitals meeting the Phase I inclusion criteria [18], and are geographically diverse, providing strong external validity within rural Texas. National and international generalizability is harder to judge, as there is likely greater variation in factors affecting quality of care.

Given the complexity of the factors influencing the effectiveness of quality improvement training programs and continuous quality improvement, and the heterogeneity of the context, content and application of such interventions, randomized trials alone may not provide adequate evaluation [22–27]. Methods providing details on contextual, environmental and implementation differences between participants who achieve positive results and those who achieve little to no improvement can elucidate influential components.

Several factors might have contributed to our lack of positive results. First, since meaningful quality improvement frequently requires substantial change in the organizational culture and workflow; 12 months follow-up may be insufficient to reap the benefits of the educational program. We plan to conduct a second evaluation at 24 months. Results may have been influenced by the inclusion of many critical access hospitals—which have less stringent reporting requirements for the quality measures examined, and so less incentive to monitor performance. The incomplete participation of approximately one-third of the intervention group was likely highly influential in the null result observed. This was partially addressed through analyses considering only those hospitals that participated in the full intervention program, or at least the majority thereof, which produced similar results (data not shown) to the intention-to-treat analysis presented here. However, the numbers in these subgroups were too small to produce conclusive results.

Other likely influential factors include the high turnover of staff at participating hospitals (affecting the continued

Table 1 Median volume of admissions eligible for each study measure during the 12-month baseline and follow-up periods for each of the study groups and subgroups analyzed in the hospital randomized controlled trial of a formal quality improvement educational program in rural Texas hospitals (October 2005 to September 2007)

<table>
<thead>
<tr>
<th>Core measures</th>
<th>Period</th>
<th>Control group (n = 23)</th>
<th>Intervention group (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ventricular function assessment</td>
<td>Baseline</td>
<td>48 (22, 80)</td>
<td>65.5 (25, 117)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>49 (17, 82)</td>
<td>73 (23, 117)</td>
</tr>
<tr>
<td>Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction</td>
<td>Baseline</td>
<td>9 (3, 16)</td>
<td>15 (10, 34)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>9 (3, 18)</td>
<td>11 (6, 30)</td>
</tr>
<tr>
<td>Heart failure composite</td>
<td>Baseline</td>
<td>48 (22, 80)</td>
<td>65.5 (25, 117)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>49 (17, 82)</td>
<td>73 (23, 117)</td>
</tr>
<tr>
<td>Oxygenation assessment</td>
<td>Baseline</td>
<td>70 (49, 122)</td>
<td>77 (51, 160)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>69.5 (41, 141)</td>
<td>83 (42, 140)</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>Baseline</td>
<td>48 (22, 92)</td>
<td>49 (19, 94)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>48.5 (29, 105)</td>
<td>49 (27, 106)</td>
</tr>
<tr>
<td>Antibiotics within 4 h</td>
<td>Baseline</td>
<td>55 (38, 96)</td>
<td>63.5 (27, 135)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>48.5 (29, 112)</td>
<td>59 (30, 131)</td>
</tr>
<tr>
<td>Pneumonia composite</td>
<td>Baseline</td>
<td>70 (49, 122)</td>
<td>77 (52, 160)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>72.5 (46, 153)</td>
<td>86 (44, 148)</td>
</tr>
<tr>
<td>All measure composite&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Baseline</td>
<td>128 (70, 219)</td>
<td>196 (68, 284)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>129 (59, 255)</td>
<td>185.5 (54, 304)</td>
</tr>
</tbody>
</table>

<sup>a</sup>All measure composite comprises the congestive heart failure and pneumonia measures above, plus five acute myocardial infarction process measures (aspirin at arrival; aspirin at discharge; angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction; beta blockers at arrival and beta blockers at discharge).
presence of individuals with continuous quality improvement training) and the mismatch between the target audience for the continuous quality improvement program and representatives sent by the hospitals. These have been discussed elsewhere [19]. Our lack of positive results may therefore speak less to the effectiveness of the educational program than to our effectiveness in reaching its target audience and/or the lack of an audience sufficiently constant and influential to effect the necessary changes. Another factor to consider is that this study examined the impact of an isolated quality improvement educational program. While the educational program is thought to be an important factor in Baylor Health Care System’s core measures performance over the past several years [17], it was not the only initiative targeting these measures. Other components not replicated here include: (i) placing a portion of administrators’ compensation at risk dependent on core measures performance [28]; (ii) a physician champion model in which paid physicians with continuous quality improvement training help lead improvement efforts and (iii) involvement of the Board of Trustees, including considerable time spent reviewing core measures performance at each monthly meeting. Based on observation in two Pennsylvania community hospitals, the ABC Baylor program can succeed outside of Baylor Health Care System when it occurs within the context of investment in other improvement strategies—such as training and funding physician champions and engaging hospital boards directly through site visits and assessments [29]. In contrast, our intervention group representatives operated with little organizational support. Despite the requirement that one physician from each participating hospital attend the educational program (incorporated in the consent and commitment form), only three hospitals had a physician to participate [19]. It should also be noted that the observed effects were smaller than the detectable effects estimated before study initiation. The estimates seemed reasonable based on improvements seen within Baylor Health Care System and the high performance on many indicators.

Table 2 Intervention group by time analysis of congestive heart failure and community-acquired pneumonia core and composite measures for rural and small community Texas hospitals participating in the hospital-randomized controlled trial of a quality improvement educational program (intention to treat analysis); October 2005 to September 2007.

<table>
<thead>
<tr>
<th>Core measures</th>
<th>Compliance % (n)</th>
<th>Odds ratio (95% CI)a</th>
<th>P-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Period</td>
<td>Education (n = 22)</td>
<td>Control (n = 23)</td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>74.8 (1641)</td>
<td>79.4 (1465)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>74.1 (1583)</td>
<td>84.1 (1444)</td>
</tr>
<tr>
<td>Congestive heart failure measures</td>
<td>Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction</td>
<td>Baseline</td>
<td>86.7 (436)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>85.2 (399)</td>
<td>86.1 (366)</td>
</tr>
<tr>
<td>Heart failure composite</td>
<td>Baseline</td>
<td>71.4 (1644)</td>
<td>74.4 (1467)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>70.6 (1584)</td>
<td>80.6 (1445)</td>
</tr>
<tr>
<td>Pneumonia measuresc</td>
<td>Oxygenation assessment</td>
<td>Baseline</td>
<td>96.6 (2234)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>99.0 (1950)</td>
<td>99.4 (2037)</td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>66.4 (1370)</td>
<td>78.1 (1317)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>76.2 (1388)</td>
<td>83.4 (1564)</td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>78.3 (1844)</td>
<td>81.0 (1692)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>82.1 (1539)</td>
<td>82.9 (1523)</td>
</tr>
<tr>
<td>Antibiotics within 4 h</td>
<td>Baseline</td>
<td>65.6 (2237)</td>
<td>71.8 (2077)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>73.8 (2141)</td>
<td>78.0 (2269)</td>
</tr>
<tr>
<td>Pneumonia composite</td>
<td>Baseline</td>
<td>68.8 (4271)</td>
<td>73.6 (3987)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td>72.4 (3968)</td>
<td>79.3 (4124)</td>
</tr>
</tbody>
</table>

aThe odds ratio (education versus control) for receiving an indicated service during the follow-up period; derived from a logistic mixed model with effects for study group membership, time, group by time, hospital (random) and hospital by time (random);

bThe P-value is associated with the group by time effect F-test in a logistic mixed model with effects for study group membership, time, group by time, hospital (random) and hospital by time (random);

cAt follow-up, 21 hospitals in the educational group and 22 hospitals in the control group reported data for pneumonia measures;

dAll measure composite comprises the congestive heart failure and pneumonia measures above, plus five acute myocardial infarction process measures (aspirin at arrival; aspirin at discharge; angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction; beta blockers at arrival and beta blockers at discharge).
achieved by some hospitals nationwide [17]. The odds ratio confidence intervals reported in Table 2 are consistent with a wide range of effects (both positive and negative).

Since this study represents the first trial of a quality improvement educational program in rural hospitals, there are no equivalent reports for comparison. Considering investigations of quality improvement educational programs in urban hospitals or ambulatory care practices, the null effects observed are not atypical. A systematic review of programs teaching quality improvement to clinicians found that, of 28 studies examining clinical outcomes, 8 reported only beneficial effects, and the subgroup of controlled studies was more likely to report mixed or null effects [15]. A second systematic review of ‘quality improvement collaborative’ interventions, which include instruction in clinical and performance improvement, similarly found null to moderately positive results with variable impact on measures [30]. This seeming lack of demonstrable effect of quality improvement training appears counterintuitive in light of its success in other high-risk industries and reports from individual health care organizations of improved quality performance [7, 8]. It is possible that in the context of a large-scale controlled trial with limited funding it is hard to achieve the necessary conditions within all the participating hospitals or practices: (i) strong commitment of upper management through direct involvement in continuous quality improvement; (ii) training management as well as frontline staff in continuous quality improvement; (iii) appreciation of small wins on a project-by-project basis and (iv) development of organizational structures and procedures to support continuous quality improvement initiatives and foster total employee involvement [31, 32]. The small number of rigorous studies of quality improvement educational programs makes it impossible to identify characteristics associated with success: for example, the kinds of clinical conditions/treatments most amenable to improvement through such approaches, the necessary mix of team members and the structure of the educational programs [33, 34].

While our study, as the first randomized controlled trial (to our knowledge) of a continuous quality improvement educational program in rural hospitals, provides valuable information regarding quality improvement and training in this setting, the lack of positive results does not conclusively show such programs are ineffective. Our results are compatible with a wide range of effects—many, clinically meaningful. Further research incorporating additional infrastructure supporting continuous quality improvement, such as physician champions, financial incentives and greater involvement of senior leadership, is needed to assess the value of quality improvement educational programs in rural hospitals.

Acknowledgments

The authors thank Briget da Graca for background research, writing and editorial assistance in preparing this article.

Conflict of interest: Baylor Health Care System provides the educational program on which the intervention in this trial is based to its employees and to external audiences, and receives compensation for the latter. No authors have any other conflicts of interest to declare.

Funding

This study is funded by the Agency for Healthcare Research and Quality (Rural Hospital Collaborative for Excellence Using IT [1UC1HS015431-01]).

References


Accepted for publication 27 March 2009