Impact of a statewide intensive care unit quality improvement initiative on hospital mortality and length of stay: retrospective comparative analysis

Allison Lipitz-Snyderman, postdoctoral fellow,[1] Donald Steinwachs, professor,[1] Dale M Needham, associate professor,[2] Elizabeth Colantuoni, assistant scientist,[3] Laura L Morlock, professor,[1] Peter J Pronovost, professor[4]

[1]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, Baltimore, MD 21205, USA

[2]Outcomes After Critical Illness and Surgery (OACIS) Group, Division of Pulmonary and Critical Care Medicine, Johns Hopkins School of Medicine, Baltimore

[3]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore

[4]Quality and Safety Research Group, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore

Correspondence to: A Lipitz-Snyderman alipitz@jhsph.edu

ABSTRACT

Objective To evaluate whether implementation of the Michigan Keystone ICU project, a comprehensive statewide quality improvement initiative focused on reduction of infections, was associated with reductions in hospital mortality and length of stay for adults aged 65 or more admitted to intensive care units.

Design Retrospective comparative study, using data from Medicare claims.

Setting Michigan and Midwest region, United States.

Population The study period (October 2001 to December 2006) spanned two years before the project was initiated to 22 months after its implementation. The study sample included hospital admissions for patients treated in 95 study hospitals in Michigan (238 937 total admissions) compared with 364 hospitals in the surrounding Midwest region (1 091 547 total admissions).

Main outcome measures Hospital mortality and length of hospital stay.

Results The overall trajectory of mortality outcomes differed significantly between the two groups upon implementation of the project (Wald test χ^2=8.73, P=0.033). Reductions in mortality were significantly greater for the study group than for the comparison group 1-12 months (odds ratio 0.83, 95% confidence interval 0.79 to 0.87 v 0.88, 0.85 to 0.90, P=0.041) and 13-22 months (0.76, 0.72 to 0.81 v 0.84, 0.81 to 0.86, P=0.007) after implementation of the project. The overall trajectory of length of stay did not differ significantly between the groups upon implementation of the project (Wald test χ^2=2.05, P=0.560). Group differences in adjusted length of stay compared with baseline did not reach significance during implementation of the project (−0.45 days, 95% confidence interval −0.62 to −0.28 v −0.35, −0.52 to −0.19) or during post-implementation months 1-12 (−0.59, −0.80 to −0.37 v −0.42, −0.59 to −0.25) and 13-22 (−0.67, −0.91 to −0.43 v −0.54, −0.72 to −0.37).

Conclusions Implementation of the Keystone ICU project was associated with a significant decrease in hospital mortality in Michigan compared with the surrounding area. The project was not, however, sufficiently powered to show a significant difference in length of stay.

INTRODUCTION

The US Institute of Medicine highlighted the serious problem of patient safety and importance of evidence based quality improvement initiatives to reduce adverse events.[1] Evidence that quality improvement initiatives intended to reduce adverse events result in a measurable impact on other important outcomes, such as mortality and length of hospital stay, is limited. Without this evidence, hospitals and healthcare payers face uncertainty about whether investment in any specific quality improvement intervention will significantly benefit patients and represent a good use of limited financial resources.

The Michigan Health and Hospital Association Keystone ICU (intensive care unit) project, developed by researchers at Johns Hopkins and undertaken by the Michigan Health and Hospital Association, about 80 of its member hospitals, and researchers at Johns Hopkins Medical Institutions, is a recent example of a successful, large scale quality improvement initiative.[2-4] The project adopted a comprehensive approach to improving patient safety that included promoting a culture of safety, improving communication between providers, and implementing evidence based practices to reduce rates of catheter related bloodstream infections and ventilator associated pneumonia. Evidence based interventions for preventing catheter related bloodstream infections were promoting handwashing, full barrier precautions, skin antisepsis with chlorhexidine, avoiding the femoral site during catheter insertion, and removing unnecessary catheters. Interventions to prevent ventilator associated pneumonia included a mechanical ventilator "bundle" consisting of use of semirecumbent positioning, daily interruption of sedation infusions, and prophylaxis for peptic ulcer disease and deep venous thrombosis.[5] The project showed that measures of culture and infection rates in the intensive care unit were substantially

improved for up to 36 months after implementing the quality improvement measures.[2 4-6]

Although adverse events, such as catheter related bloodstream infections and ventilator associated pneumonia, have been associated with increased patient morbidity, mortality, and increased length of stay,[7-12] evidence that quality improvement initiatives lead to significant reductions in these outcomes is limited. We therefore evaluated changes in statewide hospital mortality and length of stay for Medicare patients in intensive care units in Michigan hospitals compared with a group of hospitals in the surrounding area.

METHODS

This research is a retrospective, observational comparative study with concurrent controls using Medicare claims data for patients aged 65 years or more admitted to intensive care units. We evaluated whether mortality and length of stay for patients in intensive care units changed after implementation of the Keystone ICU project in Michigan hospitals versus a comparison group of hospitals located in the surrounding states in the US Midwest region.

The Keystone ICU project was initiated in October 2003 and all Michigan hospitals were invited to participate. Beginning in March 2004, multiple quality improvement interventions were sequentially implemented in participating intensive care units. Seventy seven hospitals submitted data for at least one component of this initiative. The study period spanned October 2001 to December 2006, representing a start date two years before the project was initiated and an end date coincident with the latest available Medicare data at the time of this analysis.

Although implementation dates varied across hospitals, most hospitals had implemented all interventions within one year (March 2004 to February 2005). For analysis we grouped data into six discrete periods: baseline (12 months: October 2001 to September 2002), pre-implementation of the project (12 months: October 2002 to September 2003), project initiation (5 months: October 2003 to February 2004), implementation of the interventions (12 months: March 2004 to February 2005), 1-12 months post-implementation (12 months: March 2005 to February 2006), and 13-22 months post-implementation (10 months: March 2006 to December 2006). The periods were selected a priori and correlated with phases of the research project.

Study sample and setting

The study sample included all hospital admissions for adults aged 65 or more. This group accounts for about half of all patients in intensive care units[13] and such patients are at increased risk for preventable adverse events and their subsequent negative effects on patient outcomes.[14 15] We limited analysis of hospital admissions to those that included a stay in general, surgical, medical, burns, or trauma intensive care units at an eligible hospital, defined as a hospital with 50 or more acute care beds and 200 or more total eligible

admissions to the intensive care unit during the entire five years and three months of the study period. For patients with multiple admissions during the study period, we considered each admission (the unit of analysis) a unique event.

We used an intention to treat model to define the study group. As such, the study group included admissions with a stay in an intensive care unit from all eligible hospitals in Michigan (regardless of their participation in the Keystone ICU project) plus five additional out of state hospitals affiliated with hospitals in Michigan that participated in the Keystone ICU project. Unique Medicare provider identification numbers were available for 132 Michigan hospitals; 42 of these did not meet the study's eligibility criteria, resulting in a total study group of 95 hospitals. Within the study group, about 88% of hospital admissions in this analysis came from 73 hospitals that submitted data to the Keystone ICU project.

To control for potential confounding effects of geography,[16] temporal trends in outcomes, and hospital characteristics, we used a concurrent control group for comparison. This group consisted of all hospital admissions with an intensive care unit stay from hospitals located across the other 11 states in the US Midwest region (as defined by the US Census Bureau), with hospital selection based on a random sampling methodology stratified by hospital bed size and teaching status
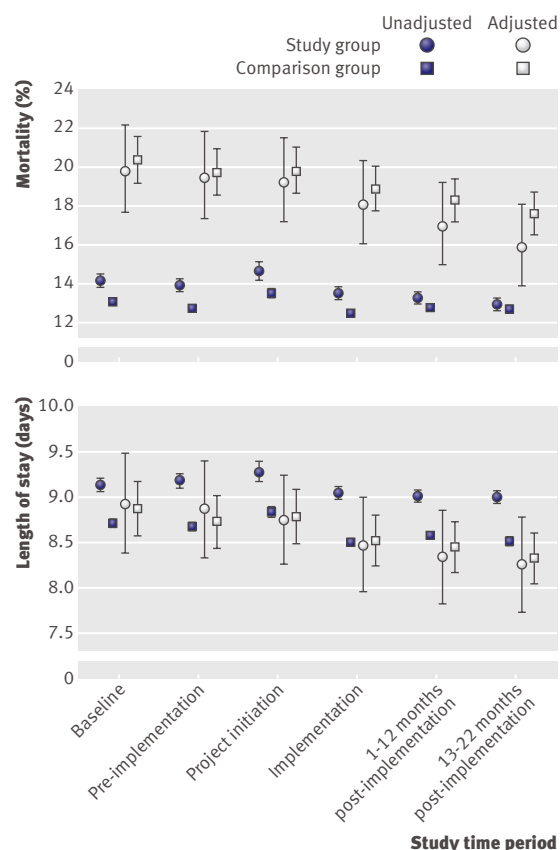
Unadjusted and adjusted population level data on hospital mortality and average length of stay in Michigan hospitals and comparison hospitals

Table 1 | Characteristics of patients in Michigan hospitals and comparison hospitals. Values are numbers (percentages)

Characteristics	Study group* (n=238 937)		Comparison group* (n=1 091 547)	
	Pre-implementation (n=110 317)	Post-implementation (n=128 620)	Pre-implementation (n=517 804)	Post-implementation (n=573 743)
Age (years):				
65-74	45 277 (41)	52 629 (41)	211 749 (41)	231 685 (40)
75-84	47 214 (43)	54 134 (42)	220971 (43)	240451 (42)
≥85	17 826 (16)	21 957 (16)	85 084 (16)	101 607 (18)
Male	54 115 (49)	63 783 (50)	260 479 (50)	287 174 (50)
White race	93 918 (85)	110744 (86)	468 576 (90)	514 413 (90)
Charlson-Deyo comorbidity†:				
Coronary heart failure	36 314 (33)	43 975 (34)	153 319 (30)	179 622 (31)
Chronic obstructive pulmonary disease	33 899 (31)	41 102 (32)	140 353 (27)	166 250 (29)
Diabetes	22 253 (20)	24 244 (19)	98 712 (19)	105 068 (18)
Acute myocardial infarction	23 019 (21)	23 718 (18)	101 918 (20)	98 404 (17)
Stroke	16 521 (15)	18 252 (14)	72 173 (14)	76 188 (13)
Primary diagnosis‡:				
Diseases of the circulatory system	54 376 (49)	56 198 (44)	258 672 (50)	254 699 (44)
Diseases of the respiratory system	14 418 (13)	17 695 (14)	59 755 (12)	71 499 (12)
Diseases of the digestive system	10 041 (9)	12 168 (9)	47 951 (9)	56 488 (10)
Injury and poisoning	8555 (8)	11 363 (9)	45 688 (9)	55 408 (10)
Neoplasms	8461 (8)	10 290 (8)	39 690 (8)	44 549 (8)
Time of discharge:				
Summer (Jun-Aug)	23 063 (21)	33 569 (26)	106 504 (21)	149 047 (26)
Autumn (Sep-Nov)	30 491 (28)	33 569 (26)	141 538 (27)	148 574 (26)
Winter (Dec-Feb)	33 724 (31)	26 111 (20)	160 224 (31)	117 237 (20)
Spring (Mar-May)	23 039 (21)	35 011 (27)	109 538 (21)	158 885 (28)

Percentages may not add to 100% owing to rounding.
*Study periods were collapsed into pre-implementation and post-implementation periods for comparison purposes. Pre-implementation include the following study periods: baseline (hospital admissions in study and comparison groups 19% and 20%, respectively), pre-implementation (19% and 20%), and project initiation (8% each); and post-implementation include the following study periods: implementation (19% each), 1-12 months post-implementation (19% each), and 13-22 months post-implementation (16% and 15%).
†Top five (out of 17) most common comorbidities are presented. Percentages do not add to 100% as comorbidities are not mutually exclusive. The remaining comorbidities have overall frequencies of ‹11%.
‡Top five (out of 18) most common clinical classification level 1 categories from Healthcare Cost and Utilization Project are presented, which represent about 86% of all patients.

(364 hospitals). Using this sampling methodology, we initially selected 64% (n=631) of hospitals in the Midwest region for the comparison group, including all large teaching, large non-teaching, and small teaching hospitals in the region. After exclusions, the final comparison group represented 37% (n=364) of Midwest hospitals. The sample size calculation was conservatively based on the number of hospitals, rather than hospital admissions, because the analysis accounted for clustering of admissions within hospitals. The sample size calculation assumed a power of 0.80, a two tailed α of 0.05, detection of a difference in reduction in length of stay between study and comparison group hospitals of 0.1 days (group standard deviation 0.3), and a fixed study group size of 90 hospitals (an initial conservative estimate of the number of Michigan hospitals during the study period), resulting in a 4:1 ratio of comparison to study group hospitals.

Data sources and study variables

For the analysis we used the dataset of the Medicare Provider Analysis and Review (MEDPAR). This dataset contains claims submitted by acute care hospitals to the Centers for Medicare and Medicaid Services for services provided to fee for service Medicare beneficiaries, with reliable identification of stays in hospitals and intensive care units. Each claim includes patient demographic data, diagnoses and procedures, an intensive care unit indicator, hospital mortality status (defined as "death" for discharge location), and length of stay. For this study we created categorical variables for patients' age (65-74, 75-84, and ≥85) and race (white and all other). We used diagnosis codes according to ICD-9-CM (international classification of diseases, ninth revision, clinical modification codes), categorised using the clinical classification software of the Healthcare Cost and Utilization Project (18 level 1 categories)[17] and included as individual indicators in the model. To adjust for patient case mix we also included individual comorbidities using the Charlson-Deyo method.[18] To account for seasonality, we used a variable identifying the three month period (quarter) in which the patient was discharged from hospital.

Hospital bed size and teaching status were obtained from American Hospital Association statistics for

Table 2 | Characteristics of Michigan hospitals and comparison hospitals. Values are numbers (percentages)

Characteristics	Study group* (n=95)	Comparison group (n=364)
Hospital bed size:		
50-199	45 (47)	153 (42)
200-299	19 (20)	72 (20)
300-399	16 (17)	60 (17)
≥400	15 (16)	79 (22)
Hospital teaching and urban status†:		
Teaching and urban	41 (43)	169 (46)
Teaching and non-urban	3 (3)	18 (5)
Non-teaching and urban	23 (24)	117 (32)
Non-teaching and non-urban	25 (26)	60 (17)
Hospital ownership:		
Non-profit	83 (87)	293 (81)
Proprietary	3 (3)	26 (7)
Government	9 (10)	45 (12)

*Five hospitals were located outside of Michigan but affiliated with hospitals in Michigan that participated in Keystone ICU project.

†Urban and non-urban location was not available for three study group hospitals (3.2%).

Percentages may not add to 100% owing to rounding.

2003. Consistent with previous publications, we modelled hospital size as a categorical variable: 50-199, 200-299, 300-399, >400 beds.[26] Hospitals were classified as teaching if residency training was approved by the Accreditation Council for Graduate Medical Education or American Osteopathic Association, or the hospital was a member of the Council of Teaching Hospitals of the Association of American Medical Colleges. We classified hospitals as urban if they were located in a metropolitan statistical area, as defined by the US Census Bureau. We combined variables for teaching and urban status into one variable comprising four categories: teaching and urban, teaching and non-urban, non-teaching and urban, and non-teaching and non-urban. Information on hospital ownership (government, proprietary, non-profit) was obtained from the Medicare cost reports.

Statistical analyses

We used multivariable logistic and Poisson (identity link) regression models to evaluate the admissions level adjusted odds of hospital mortality and adjusted mean length of stay for each of the five study periods compared with baseline, for both study and comparison groups. The models included an indicator for study versus comparison group, each of the five non-baseline periods (pre-implementation, project initiation, implementation, and post-implementation months 1-12 and 13-22), and the interactions of the group indicator with these periods. We used a Wald test to evaluate the global significance of differences between the groups in the trajectory of outcomes upon implementation of the initiative (interaction terms for group by implementation and two post-implementation periods).

All regression models were adjusted for patient variables (age, sex, race, primary diagnosis, Charlson-Deyo comorbidities, season of discharge) and hospital variables (bed size, teaching and urban status,

ownership). We used generalised estimating equations to account for clustering of hospital admissions within hospitals. In addition, the Poisson models allowed for over-dispersion by scaling the standard errors using the Pearson statistic. In our analysis of length of stay, we used the observed length of stay for all patients, including those who died during admission, as it is the observed length of stay that is of relevance to most stakeholders when considering the utilisation of resources. We obtained unadjusted and adjusted population level mortality rates and average length of stay during each study time interval for each group from the logistic and Poisson regression analyses, respectively, using the mean of predicted probabilities method.[19]

The models used in the analysis can influence results. To test whether the results were robust to alternative models, we carried out sensitivity analyses accounting for patients with multiple admissions (by including index hospital admissions only) and hospitals that closed or opened during the study period (defined as a hospital contributing no patients to the dataset during the first or last period). We also explored whether findings were sensitive to increasing degrees of patient level adjustment for case mix (from lowest to highest levels of adjustment, models included: Charlson-Deyo index and surgical and medical dichotomy for primary diagnosis, Charlson-Deyo individual comorbidities and surgical and medical dichotomy for primary diagnosis, and final model: Charlson-Deyo individual comorbidities and Healthcare Cost and Utilization Project clinical classification categories for primary diagnosis). Although we believe that hospital mortality would more closely reflect the impact of this intensive care unit based patient safety initiative, we also ran the model using 30 day mortality as the outcome variable (death dates were obtained from the Medicare denominator file). We also carried out a sensitivity analysis for length of stay excluding patients who died during admission.

SAS software, version 9.1 was used for all analyses. We considered a two sided P value <0.05 as significant.

RESULTS

Patient characteristics for hospital admissions were similar for both the study group and the comparison group (n=238 937 and n=1 091 547, respectively) and the pre-implementation and post-implementation periods (table 1). The 95 study hospitals had a higher proportion of hospitals with fewer than 200 beds and of non-teaching and non-urban status than the 364 comparison hospitals (table 2). The figure presents the population level unadjusted and adjusted mortality rates and average length of stay over the study period by group.

Adjusted hospital mortality

The overall trajectory of mortality differed significantly between the study and comparison groups upon implementation of the initiative (Wald test χ^2=8.73, P=0.033). Compared with the baseline period, mortality did not differ significantly for the study group

Table 3 | Adjusted odds ratios for mortality in Michigan hospitals and comparison hospitals

Study period	Adjusted odds ratio* (95% CI)		Study group *v* comparison group	
	Study group	Comparison group	P value†	Wald test‡
Pre-implementation	0.98 (0.94 to 1.01)	0.96 (0.95 to 0.98)	0.373	—
Project initiation	0.97 (0.92 to 1.01)	0.97 (0.94 to 0.99)	0.981	
Implementation	0.90 (0.86 to 0.93)	0.91 (0.89 to 0.93)	0.513	
Post-implementation:				χ^2=8.73, P=0.033
1-12 months	0.83 (0.79 to 0.87)	0.88 (0.85 to 0.90)	0.041	
13-22 months	0.76 (0.72 to 0.81)	0.84 (0.81 to 0.86)	0.007	

*Adjusted odds ratios compare odds of death for period of interest to each group's baseline period and were calculated using logistic regression analysis. Analyses were adjusted for age, sex, race, primary diagnosis (using Healthcare Cost and Utilization Project clinical classification level 1 categories), Charlson-Deyo comorbidities, time of discharge, hospital bed size, hospital teaching and urban status, hospital ownership, with generalised estimating equations with robust variance estimation to adjust for clustering of patient admissions within hospitals.
†P values for test of equality of study and comparison group effects were obtained from interaction terms for group and time period indicators.
‡Wald test used to test global significance of differences between study and comparison groups in trajectory of mortality upon implementation of initiative (interaction terms for group by implementation and two post-implementation periods).

versus comparison group during the subsequent periods: pre-implementation (odds ratio 0.98, 95% confidence interval 0.94 to 1.01 *v* 0.96, 0.95 to 0.98, P=0.373), project initiation (0.97, 0.92 to 1.01 *v* 0.97, 0.94 to 0.99, P=0.981), and implementation periods (0.89, 0.86 to 0.93 *v* 0.91, 0.89 to 0.93, P=0.513). Reductions in mortality were significantly greater for the study group than for the comparison group during post-implementation months 1-12 (0.83, 0.79 to 0.87 *v* 0.88, 0.85 to 0.90, P=0.041) and 13-22 (0.76, 0.72 to 0.81 *v* 0.84, 0.81 to 0.86, P=0.007; table 3).

Adjusted length of hospital stay

The difference in overall trajectory of length of stay was not significant between the study groups upon implementation of the initiative (Wald test χ^2=2.05, P=0.560). Compared with baseline, the adjusted length of stay did not differ significantly between the study group versus comparison group during the pre-implementation period (−0.06 days, 95% confidence interval −0.20 to 0.08 *v* −0.15, −0.22 to −0.08, P=0.278) and project initiation (−0.10, −0.38 to 0.02 *v* −0.10, −0.23 to 0.04, P=0.494). Although group differences in adjusted length of stay compared with baseline were greater for the study group they did not reach significance during either the implementation period (−0.45, −0.62 to −0.28 *v* −0.35, −0.52 to −0.19, P=0.429) or the post-implementation months 1-12 (−0.59, −0.80 to −0.37 *v* −0.42, −0.59 to −0.25, P=0.219) and 13-22 (−0.67, −0.91 to −0.43 *v* −0.54, −0.72 to −0.37, P=0.401; table 4). The sample size (although large) was insufficient to detect the originally projected difference of 0.1 days in reductions in length of stay between groups because the observed variation in the data was about double that proposed in the sample size calculation. Therefore, the study was underpowered to conclusively show whether length of stay differed significantly.

Sensitivity analyses

As some patients had multiple admissions shortly before death (1.9% of patients during the baseline period), estimates for hospital mortality may be conservative. However, sensitivity analyses accounting for patients with multiple recent admissions (by excluding non-index admissions within a 30 day period) or including only a single index admission per patient over the entire study period, did not change the results. Also, no changes were observed when excluding hospitals that may have opened or closed during the study period. Treatment effects for hospital mortality became more pronounced as we increased levels of adjustment for patient case mix. When 30 day mortality rather than in-hospital mortality was used as the outcome variable, results trended in the hypothesised direction for the post-implementation periods but did not reach statistical significance (Wald test χ^2=5.51, P=0.138). For both study groups, patients who died after being discharged from the hospital (within 30 days of admission) were older (> 84 years) and had more comorbidities than patients who died during their hospital stay. Excluding patients who died during their hospital stay did not the change the findings for length of stay.

DISCUSSION

Implementation of the statewide Keystone ICU project was associated with a significant reduction in hospital mortality in Michigan compared with the surrounding Midwest region. Evidence to conclusively show a significant change in hospital length of stay was insufficient, but the magnitude of the observed difference, had it been statistically significant, possibly would be considered a clinically important difference. These results may be conservative, as an intention to treat model was used and all Michigan hospitals were included in the study group despite only 77% (73 out of 95) actually participating and contributing data in the project. These findings suggest that investment in successful, large scale, robust quality improvement initiatives may not only reduce adverse events and improve quality of care but also save lives.

Improvements in adjusted hospital mortality increased over time in the study group during the two post-implementation (1-12 and 13-22 months) periods. This finding has several possible explanations. Firstly,

Table 4 | Adjusted change in hospital length of stay for Michigan hospitals and comparison hospitals

	Adjusted change in length of stay* (95% CI)		Study group *v* comparison group	
Study periods	Study group	Comparison group	P value for difference†	Wald test‡
Pre-implementation	−0.06 (−0.20 to 0.08)	−0.15 (−0.22 to −0.08)	0.278	—
Project initiation	−0.10 (−0.38 to 0.02)	−0.10 (−0.23 to 0.04)	0.494	
Implementation	−0.45 (−0.62 to −0.28)	−0.35 (−0.52 to −0.19)	0.429	
Post-implementation:				
1-12 months	−0.59 (−0.80 to −0.37)	−0.42 (−0.59 to −0.25)	0.219	χ^2=2.05, P=0.560
13-22 months	−0.67 (−0.91 to −0.43)	−0.54 (−0.72 to −0.37)	0.401	

*Change in length of stay (days) from each group's baseline period to period of interest, calculated using Poisson regression analysis. Analyses were adjusted for age, sex, race, primary diagnosis (using Healthcare Cost and Utilization Project clinical classification level 1 categories), Charlson-Deyo comorbidities (comorbidities with <15% frequency were excluded), time of discharge, hospital bed size, hospital teaching and urban status, hospital ownership, with use of generalised estimating equations for clustering of patient admissions within hospitals.
†P values obtained from interaction terms for group and time period indicators.
‡Wald test used to test global significance of differences between study and comparison groups in trajectory of length of stay upon implementation of the initiative (interaction terms for group by implementation and two post-implementation periods).

the full benefit of the multifaceted project may have taken time to realise, as was suggested by the steadily decreasing rates for catheter related bloodstream infection over the first six quarters (18 months) after implementation of the intervention.[2] Secondly, as part of the initiative included changing culture and teamwork, these elements take longer to improve and may be expected to provide a foundation for sustained efforts and other benefits, as well as create a climate that encourages continued quality improvement. Finally, building on its initial success, the Michigan Health and Hospital Association began implementing additional quality improvement initiatives during the post-implementation period that may have improved mortality. Despite these additional initiatives, a trend indicating a mortality benefit was first observed during the implementation period, supporting the impact of the Keystone ICU project on mortality.

Limitations of the study

Although this study provides important new data on the impact of a large scale quality improvement project on hospital mortality and length of stay, it has potential limitations. Firstly, these findings cannot definitively attribute the mortality benefit to this initiative, particularly given the non-randomised study design and the other subsequent efforts to improve quality and patient safety. However, no other known large scale initiatives were introduced across Michigan during the implementation period, and inclusion of a large and diverse comparison group excluded region specific temporal trends or national quality improvement efforts as an explanation for the findings. Differential trends between the study and comparison groups on changes in the use of long term acute care facilities, use of skilled nursing facilities, number of intensive care unit beds, transfers between hospitals, rates of elective surgery, or discharge practices over the study period may have contributed to the findings observed. Secondly, Medicare Provider Analysis and Review is an administrative billing database for Medicare. As a result, miscoding may occur and residual confounding may exist for differences in patient case mix. However, there are no

data to suggest large scale differences in coding or patient case mix over time or between the two groups during the study period. Furthermore, this dataset includes about half of all patients in intensive care units[13] with reliable identification of hospital and intensive care unit stay. Thirdly, generalisability of study findings may be limited to patients aged 65 or more; however, these patients represent about half of all patients in intensive care units.[13] Also, findings from this study may not be generalisable to hospitals with fewer than 50 beds.

Finally, as the project is a multifaceted quality improvement initiative with multiple interventions, the observed benefits cannot be attributed to specific components of the project. For example, the relative importance of the Comprehensive Unit Based Safety Program (the programme used to improve culture and teamwork) or the increased use of evidence based practices is unknown. However, given the low costs and risks, we believe that all components, along with robust measurement and feedback of results, may act synergistically and should be implemented together. Centralised and standardised measurement and evidence based interventions, local modification and implementation, adequate support, as well as hospital and staff readiness to engage in efforts to change, may be critical factors to yielding similar success in other settings.[20] Indeed, the combined intervention is now being implemented nationwide and in several other countries, with these factors given important consideration.

Comparison with other studies

Evaluations of quality improvement interventions intended to reduce infections often measure infection rates directly but rely on estimates from literature on cost of illness to calculate potential savings for mortality and length of stay.[21-23] However, estimates of savings vary widely owing to variations in study settings, populations, and methods for adjustment of case mix.[7-12] Since cost of illness studies compare outcomes for patients with and without infections, the extent to which adjustment for case mix sufficiently accounts for underlying differences in patient populations is

WHAT IS ALREADY KNOWN ON THIS TOPIC

Quality improvement initiatives have shown reductions in hospital acquired infections, but their impact on other important outcomes is poorly understood

WHAT THIS STUDY ADDS

Reductions in hospital mortality for Medicare patients in the intensive care unit were significant after implementation of a comprehensive statewide quality improvement initiative in Michigan

unknown. Without evidence of savings for mortality or length of stay from efforts to reduce the rate of infections, hospitals and healthcare payers may be sceptical of the potential benefits of investing in quality improvement initiatives. This study's findings are consistent with previous literature reporting excess mortality associated with catheter related bloodstream infections and ventilator associated pneumonia in the intensive care unit setting.[8 10-12]

Conclusions

Implementation of the statewide Keystone ICU project was associated with a significant decrease in hospital mortality in Michigan compared with the surrounding region. Owing to an insufficient sample size, evidence on whether reductions in length of stay observed in Michigan significantly differed from the surrounding region was inconclusive. Given the relatively low cost, the decreased rate of infections, and the lack of any known negative effects of the quality improvement initiative, these results strongly support governments', hospitals', and healthcare payers' investment in similar successful, large scale, robust quality improvement initiatives to maximise patient benefits.

1   Kohn L, Corrigan J, Donaldson M. To err is human: building a safer health system. National Academy Press, 1999.
2   Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006;355:2725-32.
3   Pronovost PJ, Berenholtz SM, Goeschel CA, Needham DM, Sexton JB, Thompson DA, et al. Creating high reliability in health care organizations. *Health Serv Res* 2006;41:1599-617.
4   Pronovost PJ, Berenholtz SM, Goeschel C, Thom I, Watson SR, Holzmueller CG, et al. Improving patient safety in intensive care units in Michigan. *J Crit Care* 2008;23:207-21.
5   Berenholtz S, Pham JC, Thompson DA, Needham D, Lubomski L, Hyzy R, et al. An intervention to reduce ventilator-associated pneumonia in the ICU. *Infect Control Hosp Epidemiol* 2010 (in press).
6   Pronovost PJ, Goeschel CA, Colantuoni E, Watson S, Lubomski LH, Berenholtz SM, et al. Sustaining reductions in catheter related bloodstream infections in Michigan intensive care units: observational study. *BMJ* 2010;340:c309.
7   Warren DK, Quadir WW, Hollenbeak CS, Elward AM, Cox MJ, Fraser VJ. Attributable cost of catheter-associated bloodstream infections among intensive care patients in a nonteaching hospital. *Crit Care Med* 2006;34:2084-9.
8   Melsen WG, Rovers MM, Bonten MJ. Ventilator-associated pneumonia and mortality: a systematic review of observational studies. *Crit Care Med* 2009;37:2709-18.
9   Pittet D, Tarara D, Wenzel RP. Nosocomial bloodstream infection in critically ill patients. Excess length of stay, extra costs, and attributable mortality. *JAMA* 1994;271:1598-601.
10 Mermel LA. Prevention of intravascular catheter-related infections. *Ann Intern Med* 2000;132:391-402.
11 Safdar N, Dezfulian C, Collard HR, Saint S. Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. *Crit Care Med* 2005;33:2184-93.
12 Soufir L, Timsit JF, Mahe C, Carlet J, Regnier B, Chevret S. Attributable morbidity and mortality of catheter-related septicemia in critically ill patients: a matched, risk-adjusted, cohort study. *Infect Control Hosp Epidemiol* 1999;20:396-401.
13 Angus DC, Shorr AF, White A, Dremsizov TT, Schmitz RJ, Kelley MA. Critical care delivery in the United States: distribution of services and compliance with Leapfrog recommendations. *Crit Care Med* 2006;34:1016-24.
14 Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-6.
15 Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *BMJ* 2000;320:741-4.
16 Knaus WA, Wagner DP, Zimmerman JE, Draper EA. Variations in mortality and length of stay in intensive care units. *Ann Intern Med* 1993;118:753-61.
17 Elixhauser A. Clinical classifications for health policy research, version 2: software and user's guide. Agency for Health Care Policy and Research, 1996.
18 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.
19 Potter L, Pasta DJ. The sums of squares are all the same—how can the LSMEANS be so different? Proceedings of the 5th Annual Western Users of SAS Software Regional Users Group Conference, 1997.
20 Bosk CL, Dixon-Woods M, Goeschel CA, Pronovost PJ. Reality check for checklists. *Lancet* 2009;374:444-5.
21 Veenstra DL, Saint S, Sullivan SD. Cost-effectiveness of antiseptic-impregnated central venous catheters for the prevention of catheter-related bloodstream infection. *JAMA* 1999;282:554-60.
22 Warren DK, Cosgrove SE, Diekema DJ, Zuccotti G, Climo MW, Bolon MK, et al. A multicenter intervention to prevent catheter-associated bloodstream infections. *Infect Control Hosp Epidemiol* 2006;27:662-9.
23 Shorr AF, O'Malley PG. Continuous subglottic suctioning for the prevention of ventilator-associated pneumonia: potential economic implications. *Chest* 2001;119:228-35.