Systematic quality improvement in healthcare: clinical performance measurement and registry-based feedback
van der Veer, S.N.

Citation for published version (APA):
van der Veer, S. N. (2012). Systematic quality improvement in healthcare: clinical performance measurement and registry-based feedback

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 7

The effect of a multifaceted feedback strategy on ICU patient outcomes compared to registry-based feedback reports alone

Effect of a tailored multifaceted performance feedback intervention on length of stay compared to feedback reports alone: a cluster randomized trial in intensive care.
Submitted for publication
Abstract

Context
Intensive care units (ICUs) worldwide invest valuable resources in national registries to continuously monitor clinical performance based on periodical feedback reports. Studies in other medical domains showed that feedback is more likely to affect healthcare when complemented with other strategies.

Objective
To assess the impact of a multifaceted performance feedback intervention on ICU length of stay (ICU LOS) compared to only sending standard registry benchmark reports.

Design, Setting, and Participants
The InFoQI (Information Feedback on Quality Indicators) study was a cluster randomized trial from February 2009 to May 2011 among 30 Dutch closed-format ICUs that participated in the national registry; all completed the study. Study duration per ICU was sixteen months. Cardiac surgery admissions were excluded. Finally, we analyzed data on 26077 admissions.

Intervention
Intervention ICUs received more frequent and comprehensive feedback reports, established a local, multidisciplinary quality improvement team, and received two educational outreach visits. Control ICUs received standard benchmark reports.

Outcome Measures and results
The extent to which the InFoQI program was implemented in daily practice varied considerably between intervention ICUs. The program had no statistically significant impact on any of the patient outcome measures. Our primary endpoint, ICU LOS, reduced by 3% in intervention units compared to controls (95% confidence interval [CI] -14% to +6%). Regarding the secondary measures, duration of mechanical ventilation increased by 4% (95% CI, -18% to +22%), out-of-range glucose measurements decreased 11% (95% CI, -33% to +19%), and all-cause hospital mortality decreased by 4% (95% CI, -25% to +22%).

Conclusions
Multifaceted performance feedback did not lead to better patient outcomes than only sending feedback reports. Our study underlines the difficulty of further improving ICU performance with this type of intervention within the context of well-organized healthcare systems, where levels of care are already high.

Trial Registration
Current Controlled Trials ISRCTN50542146 www.controlled-trials.com
Introduction

The intensive care unit (ICU) provides complex multidisciplinary and expensive care to a heterogeneous patient population with relatively high mortality and morbidity rates. Performance monitoring and systematic quality improvement (QI) have become increasingly common tools in the field of intensive care medicine.1-3 These tools rely on indicator sets4-9 and collection of indicator data by national registries,10-14 and they require substantial investments of scarce healthcare resources. To facilitate local QI activities, ICUs participating in national registries receive periodical feedback reports on a broad range of performance indicators, benchmarked against their own historical performance or that of other units. The underlying assumption is that inferior or inconsistent care presented in these reports prompts providers to change their practice.15

Two systematic reviews showed that in order to increase the impact of feedback reports on healthcare quality, they should be complemented with other strategies.16,17 However, of the total of 130 randomized studies included in16,17, only one pertained to adult intensive care18. Since contextual factors are known to influence the success of QI interventions,19 it was not self-evident that the results of these reviews could be extrapolated to the intensive care setting. This indicated a paucity in high-level evidence on the effect of multifaceted feedback interventions on the quality of ICU care. More recently, three randomized studies showed that such interventions can positively affect ICU practice.20-22

In order to augment this body of knowledge, we developed a multifaceted performance feedback strategy including benchmark reports on a wide spectrum of quality indicators13 as collected by the Dutch national ICU registry.13 The resulting program was named InFoQI (Information Feedback on Quality Indicators).23 Although the context of InFoQI resembles that of ICUs worldwide receiving registry benchmark reports to monitor and improve their care, we are not aware of previous randomized studies evaluating a multifaceted feedback intervention within this specific registry context. If the program would be successful, those receiving and providing registry benchmark reports should consider complementing the reports with additional strategies to effectively support and accelerate systematic, local QI at ICUs.

The aim of the current study was to evaluate in intensive care the impact of the InFoQI program on ICU length of stay (ICU LOS), mechanical ventilation duration, mortality, and glucose regulation after one year of receiving the intervention compared to only sending quarterly standard benchmark reports. Therefore, we conducted a cluster randomized trial among ICUs in the Netherlands.

Methods

The Dutch National Intensive Care Evaluation registry

The Dutch National Intensive Care Evaluation (NICE) registry aims to systematically and continuously monitor and improve ICU performance by reporting and benchmarking quality indicators. They started in 1996 with the outcome indicators case-mix adjusted hospital mortality and ICU LOS,13 at the time of the current study, a sample of 80 ICUs –covering 85% of all Dutch ICUs– voluntarily submitted these core data to the NICE registry. Recently, the core set was extended to a total of eleven structure-, process- and outcome indicators.6

Regular NICE services include standard quarterly and annual benchmark reports on the core set, complemented with similar, but separate reports on the extended indicator set. Also, the NICE team of data managers/researchers and software engineers supports registry participants with additional data analyses.
Study design
We randomized ICUs (i.e., clusters), because the intervention was targeted at the facility rather than patient level.24 Our study was also pragmatic.25 A detailed description of our design was published elsewhere.23

Participating ICUs
All ICUs in the Netherlands are closed-format,26 and the large majority has an intensivist on call around the clock. We regarded the intensivists being responsible for the clinical process a facilitator for systematic, local QI activities.27 ICUs were eligible for the InFoQI study if they participated in the NICE registry and were preparing to submit data to the registry on the extended indicator set. They had to be able to allocate at least two staff members for a minimum of four hours per month for InFoQI activities.

All patients admitted to the participating ICUs during the study period were included in the analysis, except for admissions following cardiac surgery, and patients admitted to prepare for organ donation.

Multifaceted performance feedback intervention
For ICUs assigned to the intervention arm, the regular NICE services were extended with:
- twelve monthly reports focusing on monitoring local performance over time, and four comprehensive quarterly benchmark reports on the extended indicator set facilitating comparison with other ICUs;
- establishment of a local, multidisciplinary QI team that had to consist of at least one intensivist and one nurse. The team’s main tasks included formulating a QI action plan, monthly monitoring and discussing of performance using the feedback reports, and initiating QI activities.
- two educational outreach visits by the investigators to support the QI team with interpreting performance data, and identifying opportunities for improvement.

Feedback reports based on indicator data combined with an educational component, and the development of a QI plan had been reported to potentially improve care.28 We further tailored the intervention to prospectively identified barriers29,30 to using performance data for QI activities, e.g., lack of trust in data quality, and having difficulties to interpret the feedback. A detailed description of the intervention and identified barriers was published elsewhere.23

Units allocated to the control arm received regular NICE services.

Outcome measures
We selected the quality indicator ICU LOS as the primary endpoint of our study. Firstly, because we expected successful QI actions aimed at other indicators to contribute to an improvement of ICU LOS. Secondly, NICE data from 2008 showed that ICU LOS was the indicator showing the largest variation among ICUs, when corrected for admission type. The indicators mechanical ventilation duration, proportion of glucose measurements outside the range of 40 to 144 mg/dl, and all-cause hospital mortality were selected as secondary endpoints.

Cluster randomization and allocation
Randomization of ICUs was stratified by size (more/less than the national median number of ventilated, non-cardiac surgery admissions) and involvement (yes/no) in a previous indicator development pilot to evaluate feasibility of data collection.6 Per stratum, we generated a randomization scheme with variable block sizes using dedicated software. This scheme was
concealed to those enrolling and assigning ICUs. Due to the character of the intervention, it was not possible to blind participants or those involved in providing the InFoQI program.

**Data collection and validation**
We used the available information infrastructure of the NICE registry, in which participants either manually entered data using dedicated software, or automatically extracted data from electronic patient records. They uploaded their local data monthly to the central NICE registry database. The registry’s infrastructure routinely provides data quality assurance.

To evaluate the extent to which the InFoQI program was implemented as planned we asked individual QI team members twice during the study period to record their activities, including the estimated time they invested.

**Statistical analysis**
The total study period for intervention ICUs lasted sixteen months, starting at randomization and ending three months after the last report was sent. The period between randomization and the first outreach visit was approximately two months and marked as pre-InFoQI, directly followed by the InFoQI period. The pre-InFoQI period for control ICUs was defined as the first two months after randomization, followed by a fixed InFoQI period of fourteen months (Figure 1).

In all analyses, we tested for the effect of arm (intervention versus control), time since start InFoQI period (with value ‘0’ for all admissions during the pre-InFoQI period), and the interaction between arm and time. We focused on the interaction term to assess the difference in change at the end of the InFoQI period between the two arms, because we expected intervention ICUs to improve gradually.

---

**Figure 1: The pre-InFoQI and InFoQI period for intervention and control ICUs.**
The duration of the pre-InFoQI and InFoQI period in the intervention arm varied between units because this depended on, for example, how soon the first outreach visit could be scheduled so that all QI team members were able to attend. The months in the upper panel, therefore, reflect the mean values of all intervention ICUs. For units in the control arm the duration of both periods was fixed.
For ICU LOS we performed a Cox proportional hazard regression analysis to the subdistribution hazard\(^33\) of the time to ICU discharge, with ICU death as competing risk.\(^34,35\) The length of stay of the first ICU admission was prolonged with the length of stay of subsequent ICU readmissions within the same hospital admission. Furthermore, we analyzed the time to ICU death, with ICU discharge as the competing risk. For duration of mechanical ventilation, we applied the same procedure analyzing the time to extubation, with death within six hours after extubation as the competing event. To analyze the proportion of out-of-range glucose measurements we used binomial regression with a logistic link function, including only admissions with an ICU LOS >72 hours because we expected the benefit of improved glucose regulation to be most pronounced in this group.\(^36\) Logistic regression analyses were used to verify that the intervention did not increase all-cause hospital mortality or readmission rates.

To adjust for differences in case mix between the study arms, we used four patient-level variables (age; sex; Acute Physiology and Chronic Health Evaluation (APACHE) IV score; admission type) and two ICU-level variables (academic/teaching or non-teaching unit; participation in indicator development pilot) as covariates in each regression analysis. We used natural splines to model non-linear effects of continuous variables (age, APACHE IV score). We chose a marginal modeling approach in all analyses to account for potential correlation of outcomes within ICUs;\(^37\) in the Cox regression we followed the method of Lin and Wei,\(^38\) while in the binomial and logistic regression analyses we used generalized estimating equations, with exchangeable working correlation.\(^39\)

To further explore the impact of our intervention, we conducted a post-hoc as-treated analysis of ICU LOS, comparing all ICUs from the original control arm to the intervention ICUs that reported a monthly time investment exceeding four hours per QI team member. Our sample size calculation\(^23\) showed that we needed a sample of 26 ICUs to detect a relative reduction in ICU LOS of 27%—corresponding to an absolute reduction of 0.5 days—with 80% power at a type I error risk of 5%, taking an estimated intra-cluster correlation of 0.036 into account.

We used R version 2.13.1 for statistical analyses.

**Results**

**Participants**

Of the 80 ICUs submitting core data to NICE, 46 were preparing data collection on the extended set; 30 accepted our invitation to participate in the trial (Figure 2). The main reason to refuse was a lack of resources to establish a local QI team (n=7). Fifteen units were assigned to the intervention arm, and an equal number to the control arm; all completed the study. ICUs were enrolled in the study between February and December 2009. All were mixed medical-surgical units.

In total, there were 35196 admissions during the study period. We excluded 4996 admissions following cardiac surgery (14.2%), and 24 admissions for organ donation (<0.1%). We also excluded 3460 admissions for which—according to the APACHE IV criteria—we could not calculate a severity of illness score (9.8%), and 639 admissions with one of the other case mix variables missing (1.8%). Finally, we included 30 ICUs and 26077 admissions in our analysis. Table 1 displays the baseline characteristics of both arms at the level of ICUs and admissions.

For glucose regulation, four intervention ICUs failed to submit data due to technical problems with their local laboratory system interface, and were excluded from this part of the analysis.
Chapter 7

Figure 2 Flow diagram of ICUs and patients through the trial

Total ICUs in Netherlands (n=94)

Total ICUs participating in NICE registry (n=80)

Total ICUs preparing to submit data on the extended quality indicator set (n=46)

Reasons to refuse (n=16)
• Insufficient resources to form QI team (n=7)
• No commitment to participate (n=4)
• Unable to timely submit data (n=3)
• Involved in reorganization (n=2)

Randomized (n=30)

15 ICUs allocated to intervention arm; all received allocated intervention (median [range] admissions per ICU, 725 [530-2014]; total numbers of admissions, 17442)

15 ICUs allocated to control arm; all received allocated intervention (median [range] admissions per ICU, 1045 [675-1615]; total numbers of admissions, 17754)

No ICUs lost to follow up

15 ICUs included in analyses (13800 total admissions)
Admissions excluded from analyses
2037 cardiac surgery
11 patients admitted for organ donation
1427 no APACHE IV score calculated
167 missing casemix variables

15 ICUs included in analyses (12277 total admissions)
Admissions excluded from analyses
2959 cardiac surgery
13 patients admitted for organ donation
2033 no APACHE IV score calculated
472 missing casemix variables
### Table 1 Baseline characteristics of participating ICUs (clusters) and admissions

<table>
<thead>
<tr>
<th></th>
<th>Intervention</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ICU level characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. included in analysis</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Median (IQR) number of admissions</td>
<td>725 (530-2014)</td>
<td>1045 (675-1615)</td>
</tr>
<tr>
<td>Academic or teaching hospital</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>&gt; 150 ventilated, non-surgical admissions per year</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Participated in indicator development pilot</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td><strong>Admission level characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. included in analysis</td>
<td>13800</td>
<td>12277</td>
</tr>
<tr>
<td>Mean (SD) age (years)</td>
<td>61.2 (16.8)</td>
<td>62.3 (17.1)</td>
</tr>
<tr>
<td>Male sex</td>
<td>7763 (56.3)</td>
<td>6856 (55.8)</td>
</tr>
<tr>
<td>Admission type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>medical</td>
<td>6573 (47.6)</td>
<td>6167 (50.2)</td>
</tr>
<tr>
<td>elective surgery</td>
<td>5177 (37.5)</td>
<td>3641 (29.7)</td>
</tr>
<tr>
<td>emergency surgery</td>
<td>2050 (14.9)</td>
<td>2469 (20.1)</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td>5745 (41.6)</td>
<td>5791 (47.2)</td>
</tr>
<tr>
<td>Mean (SD) APACHE IV score</td>
<td>57.4 (31.5)</td>
<td>59.8 (33.2)</td>
</tr>
</tbody>
</table>

Abbreviations: APACHE, Acute Physiology and Chronic Health Evaluation; CI, Confidence Interval; ICU, Intensive Care Unit; IQR, Interquartile range; QI, Quality improvement; SD, Standard deviation.

Values are numbers (percentages), unless indicated otherwise.

a) Previous pilot to evaluate the feasibility of indicator data collection

---

**Implementation of the InFoQI program in daily practice**

All fifteen intervention units established a local QI team in which intensivists and ICU nurses were represented; thirteen ICUs complemented this team with other representatives, e.g. an operational manager. All ICUs received both educational outreach visits. The average monthly time investment per member was 4.1 hours (standard deviation [SD], 2.3; range, 0.6 to 8.1). As planned, all intervention ICUs received 4 quarterly and 12 monthly reports. The average number of reports reviewed by at least one team member was 10.6 (SD, 2.8; range, 3 to 16), while the average number of monthly QI team meetings to discuss the reports was 5.7 (SD, 1.4; range, 0 to 12). For units that spent at least 4 hours per month per team member (n=8) this was 13.2 reports (SD, 2.8; range, 8 to 16), and 9.1 meetings (SD, 2.3; range, 5 to 12). None of the ICUs were able to review and discuss all reports.

The QI action plans formulated during the outreach visits consisted of a mean of 12.2 planned actions (SD, 3.5; range, 6-17). Of all quality indicators, glucose regulation was the most actionable with an average of 2.5 actions (SD, 1.5; range, 1 to 5). Table 2 contains the type of actions and examples for each outcome measure.

**Effect of the intervention**

Table 3 shows that ICUs that received the multifaceted feedback intervention did not improve their patient outcome measures more than ICUs only receiving standard benchmark reports. ICU LOS reduced 3% after one year in intervention units compared to controls (95% confidence interval [CI], -14% to +6%) (Figure 3). This reduction supplemented an already existing, non-significant difference in ICU LOS in the pre-InFoQI period between intervention and control units of 9% (95% CI, -29% to +8%).

---

167
<table>
<thead>
<tr>
<th>Type of action</th>
<th>ICU LOS</th>
<th>Duration of mechanical ventilation</th>
<th>Out-of-range glucose measurements</th>
<th>Hospital mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigate low performance for patient subgroups</td>
<td>Investigate effect of delayed ICU discharge of surgery patients due to unavailability of ward beds</td>
<td>Investigate effect of case mix of ventilated surgical patients using a group of similar ICUs as comparison</td>
<td>Analyze high % hyperglycemia in patients admitted after 'surgery for cranial neoplasm'</td>
<td>Analyze standardized mortality ratio &gt;1 for medical admissions</td>
</tr>
<tr>
<td>Investigate individual cases within QI team</td>
<td>Discuss patients who were admitted / ventilated longer than the national 90th percentile</td>
<td>Review episodes of hypoglycemia</td>
<td>Investigate records of patients who died despite a low mortality risk</td>
<td></td>
</tr>
<tr>
<td>Share QI team findings in ICU staff meetings</td>
<td>Organize 6-weekly multidisciplinary meetings to discuss individual admissions with long ICU LOS / duration of mechanical ventilation</td>
<td>Organize monthly meetings with ICU nurses to discuss causes of episodes of hypoglycemia</td>
<td>Organize monthly mortality conferences with intensivists and anesthesiologists</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>[No actions formulated]</td>
<td>Organize sessions to refresh knowledge on sedation protocol</td>
<td>Formalize instructions on glucose regulation for temporary ICU nurses</td>
<td>[No actions formulated]</td>
</tr>
<tr>
<td>Adjust protocols or care processes</td>
<td>Appoint lead nurse for admissions with expected ICU LOS &gt;7 days</td>
<td>Develop wean protocol with regards to regulation of values exceeding 216 mg/dl</td>
<td>[No actions formulated]</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ICU, Intensive Care Unit; LOS, Length of stay; QI, Quality improvement.

a) value below 40 mg/dl

b) value exceeding 144 mg/dl

c) includes newly established as well as existing meetings
Effect of registry-based feedback on ICU patient outcomes

Figure 3: Time to alive discharge from the ICU for both arms after receiving the intervention for one year

The time to ICU death was not affected by the intervention (-3%; 95% CI, -40% to +25%). The number of hospital deaths in the intervention arm compared to controls decreased 4% after an InFoQI period of one year (95% CI, -25% to +22%), and the proportion of glucose measurements outside the range of 40 to 144 mg/dl decreased 11% (95% CI, -33% to +19%). The duration of mechanical ventilation increased 4% in intervention ICUs compared to controls (95% CI, -18% to +22%). InFoQI did not significantly affect the readmission rate (-13%; 95% CI, -34% to +13%).

The results of the post-hoc as-treated analysis resembled those of our primary analysis: after receiving the intervention for one year, ICU LOS in as-treated units reduced by 4% compared to controls, but without reaching statistical significance (95% CI, -14% to +5%).

Discussion

We randomized 30 closed-format ICUs that participated in the Dutch national registry, and analyzed data on over 26,000 admissions to evaluate the effect on patient outcome measures of a multifaceted performance feedback program—including local, multidisciplinary QI teams, and educational outreach visits—compared to only sending standard benchmark reports. The extent to which the InFoQI program was implemented in daily practice varied substantially between ICUs in the intervention arm. The program had no impact on ICU LOS, or on any of the secondary outcome measures.

Strengths and limitations

A strength of our study was that we built on the established infrastructure of the Dutch national registry which enabled ICUs participating in InFoQI to rely on routine registry procedures without requiring any additional data collection activities. This decreased the risk of
demotivated control ICUs –potentially leading to a Hawthorne effect or control units discontinuing their participation due to an increased workload without the advantage of receiving the intervention. Also, the registry’s quality assurance framework accounted for all recommended data quality control methods for evaluating QI programs, which increased the completeness and reliability of our data.

The cluster randomized design is another strength of our study, because it helped avoiding bias in the estimates of ICU performance resulting from confounding with known and unknown unit-level characteristics. Despite stratified randomization, more intervention than control ICUs participated in the indicator development pilot. To correct for the potential influence of this baseline imbalance, we added this as a covariate in our analyses. Furthermore, our analysis strategy prevented us from falsely interpreting a decrease in ICU LOS as a positive effect of the intervention, while in fact this may have been caused by more ICU deaths, and more premature discharges leading to readmissions.

We analyzed data on over 26000 admissions, which –to our knowledge– is the largest number of patients included in a randomized trial evaluating a QI strategy in intensive care so far. Although the adjustment for potential correlation of outcomes within ICUs decreased the effective sample size, repeating the power analysis using trial data showed that 23 centers would have been sufficient to detect the reduction in ICU LOS that we considered relevant beforehand. Also, the actual intra-cluster correlation coefficient of 0.021 was lower than anticipated. This implies that the study was not underpowered.

One limitation was that the InFoQI program aimed to intervene at the organizational level, where it might take longer to effectuate change than at the individual physician or patient level. For example, an intervention ICU that identified the hospital ward responsible for the majority of the delayed discharges, and thus for increasing ICU LOS, was still working on a solution with the ward’s management at the end of the study period. Hence, it is possible that prolonging the intervention and follow-up period would have increased the probability of finding a statistically significant effect of the program.

Another limitation is that we had mainly outcome of care measure available as the basis for our feedback. Indicators like hospital mortality and ICU LOS are influenced by several providers and practices, and factors other than ICU care. This impedes feedback interpretation, assignment of accountability, and identification of effective actions. By contrast, process of care measures directly assess provider action, making them more actionable. This was confirmed by our finding that the proportion of out-of-range glucose measurements was the indicator with most QI actions.

Relation to other studies
Other cluster randomized trials within the ICU domain evaluated the effect of a multifaceted QI intervention including performance feedback. In contrast to our study, the control ICUs in these trials received no feedback, increasing the difference between both arms. Moreover, they focused on improving a specific part of ICU care, e.g. by preventing accidental extubations. All were successful in increasing adherence to best practice, but only few also evaluated the effect of their intervention on outcome of care measures. One study that improved adherence to feeding guidelines, reported a non-significant impact on ICU LOS. Likewise, a non-randomized study that reported a decrease in catheter-related bloodstream infections and a reduction in hospital mortality in ICU patients, did not show a significant difference in hospital LOS. This illustrates the difficulty to define process of care measures with a sufficiently strong link to length of stay.
<table>
<thead>
<tr>
<th>No of clusters / admissions included in analysis</th>
<th>Crude outcome (median (IQR), unless indicated otherwise)</th>
<th>Crude difference in change over time between arms</th>
<th>Adjusted (^a) difference in change over time between arms</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Control</td>
<td>Intervention</td>
<td>Control</td>
<td>Main effect (95% CI)</td>
</tr>
<tr>
<td>ICU Length of stay (days)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 / 13800</td>
<td>15 / 12277</td>
<td>1.1 (0.8-3.5)</td>
<td>1.5 (0.8-4.2)</td>
<td>1.03 (0.96-1.11) (^b)</td>
</tr>
<tr>
<td>Duration of mechanical ventilation (days)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 / 5745</td>
<td>15 / 5791</td>
<td>1.6 (0.4-5.7)</td>
<td>1.8 (0.5-6.0)</td>
<td>0.94 (0.84-1.04) (^b)</td>
</tr>
<tr>
<td>Proportion of out-of-range (^c) glucose measurements (^d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 / 2608</td>
<td>15 / 3807</td>
<td>0.29 (0.15-0.46)</td>
<td>0.30 (0.17-0.45)</td>
<td>0.86 (0.64-1.15) (^e)</td>
</tr>
<tr>
<td>All-cause hospital mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 / 13800</td>
<td>15 / 12277</td>
<td>2021 (14.6) (^f)</td>
<td>2128 (17.3) (^f)</td>
<td>0.95 (0.80-1.14) (^e)</td>
</tr>
</tbody>
</table>

Abbreviations: CI, Confidence Interval; ICC, Intracluster correlation coefficient; ICU, Intensive Care Unit; IQR, Interquartile range; P, probability.

\(^a\) Adjusted for age, sex, APACHE IV score, admission type, academic/teaching or non-teaching unit, and participation in previous indicator development pilot

\(^b\) Hazard ratio for the interaction term between arm and time, reflecting the difference in change between the two arms after one year of exposure to the intervention; a value >1 means that patients in the intervention arm reached the event of interest (e.g. alive ICU discharge) sooner than controls (i.e., had a shorter ICU length of stay).

\(^c\) Values below 40 mg/dl or exceeding 144 mg/dl.

\(^d\) Only admissions with ICU length of stay > 72 hours were originally included in the analysis (n=7617); of these, we excluded admissions (number, percentage) (i) with missing glucose measurements due to technical problems with the automated laboratory system interface (1148, 15.1); this implied excluding four intervention ICUs from the analysis, and (ii) without any glucose measurements (54, 0.7).

\(^e\) Odds ratio for the interaction term between arm and time, reflecting the difference in change between the two arms after one year of exposure to the intervention.

\(^f\) Values are numbers (percentage)
Meaning of the study
All ICUs in our study were Dutch closed-format units, with an infrastructure to routinely submit data to a national registry, and sufficient, motivated staff and supportive management to establish a QI team. Although this optimized the environment for successful implementation of the InFoQI program, the feasibility of the intervention in daily practice appeared not to be self-evident. A possible explanation is an underestimation of the required time investment for team members resulting in an unanticipated lack of local resources to perform study activities and execute the QI action plan.

Although ICUs that invested more time achieved a more complete implementation of InFoQI, our as-treated analysis suggested that this is not the only ingredient for success. We expect opportunities to lie in providing teams with additional tools to translate feedback into potentially more effective, evidence-based QI actions. For example, cause-and-effect diagrams for systematic problem analysis, or evidence-based input on how to change daily practice in order to improve performance, e.g., using a daily goals form during patient care rounds.

The generalizability of our findings is limited to high-level healthcare systems, in which a national registry has been available for some time. These contextual factors are likely to have contributed to ICU patient outcomes steadily improving over the last decade, increasing levels of care, and possibly causing a ceiling effect. Yet, areas of intensive care that are optimized in the Dutch situation might still show room for improvement elsewhere. Our intervention could, therefore, be effective in the context of less well organized systems, or in countries that only recently established a national registry to monitor ICU performance. However, in such contexts we expect the issue of feasibility to be even more tenacious.

Our study underlines the difficulty of showing benefits to patient outcomes even with motivated participants and an intervention founded on an extensive barrier analysis. Based on our results, those receiving and providing registry reports should only consider adding local QI teams and outreach visits if they have data available on actionable process measures linked to patient outcome, and are able to provide additional tools to support the translation of feedback into effective actions.

Future research
We tailored the InFoQI program to overcome barriers to using performance feedback for local QI activities. In a future qualitative study, we will evaluate which prospectively identified barriers remained untargeted, and if any other factors affected the impact, in order to find a more detailed explanation for the program’s ineffectiveness.

Future research might aim to identify actionable process of care measures. Besides a solid evidence-base link between the measures and ICU patient outcome, routine collection of reliable data on the process measures should be feasible.

Lastly, more knowledge is needed on effective and useful tools to facilitate ICU clinicians to translate performance feedback into QI actions.
Reference List

(15) Kiefe CI, Allison JJ, Williams OD, Person SD, Weaver MT, Weissman NW. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. JAMA 2001; 285:2871-2879.


