Evaluation of a critical care outreach service in a middle-income country: A stepped wedge cluster randomized trial and nested qualitative study

Alireza Jedidian, PhD a, Karla Hemming, PhD b,a, Antje Lindenmeyer, PhD b, Arash Rashidian, PhD c, Leila Sayadi, PhD a, Nazila Jafari, MD a, Reza Malekzadeh, PhD a, Tom Marshall, PhD b

a Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran
b Institute of Applied Health Research, University of Birmingham, Birmingham B15 2TT, UK
c School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Keywords: Critical care outreach, Stepped wedge randomized controlled trial, Hospital mortality, Cardiopulmonary resuscitation, Length of stay, Health service evaluation

A R T I C L E   I N F O

Keywords:
- Critical care outreach
- Stepped wedge randomized controlled trial
- Hospital mortality
- Cardiopulmonary resuscitation
- Length of stay
- Health service evaluation

A B S T R A C T

Purpose: This trial evaluates implementation of critical care outreach in a middle-income country.

Materials and methods: Critical care outreach delivered by a team of intensive care nurses was implemented across general hospital wards in an Iranian university hospital. The order of implementation was randomized with wards stratified by predicted mortality rates. Effectiveness was evaluated using a stepped wedge cluster randomized controlled trial design, comparing outcomes between patients admitted before and after implementation. The primary outcomes were inhospital mortality and cardiopulmonary resuscitation. A nested qualitative study explored challenges to implementation and contextualized the trial outcomes.

Results: Between July 2010 and December 2011, 13 wards were sequentially randomized to implement the critical care outreach: 7802 patients were admitted before implementation and 10 880 after implementation. There were 370 deaths (4.74%) among patients admitted before implementation and 384 deaths (3.53%) after implementation. Adjusting for clustering and temporal trends, the odds ratio for mortality was 1.03 (95% confidence interval, 0.68–1.53). Results for other outcomes were broadly similar. Focus groups revealed a lack of endorsement of the intervention by management and ward nurses.

Conclusions: This pragmatic evaluation of critical care outreach in a middle-income country did not show a reduction in mortality or other outcomes.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Scientific background

Demand for intensive care beds is increasing in lower income and middle-income countries [1,2]. Critical care outreach, comprising a system for identifying acutely ill patients in general wards and an outreach team, is widely implemented in developed countries [3–7]. However, systematic reviews of randomized controlled trials have not found robust evidence that it reduces mortality, cardiac arrest, unplanned intensive care admissions, or length of stay [8–10]. It has been suggested that the policy was not evidence based [11,12]. Apart from one before-and-after study, it is unevaluated in middle-income countries [13].

1.2. Explanation of rationale

Hospital managers decided to implement critical care outreach (CCO) across the general hospital wards of Shariati Hospital, Tehran. They agreed to a randomized roll-out, allowing robust evaluation as a stepped wedge cluster randomized controlled trial [14].

http://dx.doi.org/10.1016/j.jcc.2016.07.018
0883-9441/© 2016 Elsevier Inc. All rights reserved.
1.3. Aim

This trial assessed the effects of CCO on hospital mortality and cardiopulmonary resuscitation. Secondary aims were to assess effects on length of stay and intensive care admissions.

2. Methods

Between July 2010 and December 2011, Shariati Hospital implemented and sequentially randomized CCO across 13 wards as an unblinded stepped wedge cluster randomized trial. Outcomes were compared between admissions before (unexposed) and after each ward implemented CCO (exposed).

2.1. Trial design

The trial was implemented in periods of 4 weeks: baseline data collection for 3 periods (12 weeks), roll-out of the intervention to 2 wards every 2 periods (6 steps of 8 weeks each), and postintervention data collection for 3 periods (12 weeks). This was a total of 18 periods (72 weeks) (Supplementary Fig. 1). Each ward also had 8-week transition phase of implementation, during which ward staff were trained to adopting the intervention.

2.2. Rationale for the trial design

Randomization was at the cluster level to avoid issues of contamination. Because it was necessary to implement CCO sequentially in wards rather than introduce it to all wards at the same time, we randomized the roll-out sequence. This allowed us to evaluate implementation as a stepped wedge cluster randomized trial.

2.3. Participants and setting

Shariati Hospital is a university and public teaching hospital with 800 beds, in 29 wards including 5 intensive care units (47 beds). It admits 20,000 patients annually. All 13 adult general wards (general medical wards, orthopedics, hematology, obstetrics, pulmonary, urology, surgery, and maxillofacial wards) served by 3 of the 5 intensive care units were selected for the new CCO team.

There were no patient exclusion criteria, everyone admitted to the 13 wards over the duration of the trial was classified as belonging to 1 of the 3 exposure groups (unexposed, transition phase, and exposed). Those admitted before the ward was randomized to implement the intervention were unexposed, those admitted after were exposed, and those admitted when the ward was undergoing training were in the transition phase.

2.4. Intervention

Critical care outreach was intended to respond to the needs of acutely ill patients and to share skills between intensive care and general ward staff. Implementation was overseen by a committee including representatives of management, nursing, and medical teams. The CCO team included 6 experienced intensive care nurses who before the trial were introduced to the ward staff and underwent 3 months of additional training in patient monitoring and clinical management (Supplementary Appendix 1). Training of the critical care team included theory and management protocols followed by full-time practical training. The week the ward crossed over to the intervention, ward nurses began 8 weeks of training on assessment, identification, and management of acutely ill patients (Supplementary Appendix 2).

The committee chose a single parameter system using routinely measured vital signs for ward staff to use to identify acutely ill patients for the CCO team. This was simple, avoided calculations, and minimized false alerts [15]. Eligibility criteria included physiological criteria listed in Supplementary Appendix 3 (respiratory rate, oxygen saturation, pulse, blood pressure, temperature, urinary output, and change in consciousness), ward staff concern, recent discharge from intensive care, or patients actively identified by the CCO team. Eligible patients showing no improvement after 30 minutes were referred to the CCO team. The CCO team assessed these patients using a composite scoring system (Supplementary Appendix 4). The CCO team managed all high-risk patients (score $>5$) and determined who should care for moderate-risk patients (score 3-5). Ward staff managed all low-risk patients (score $<3$). Patients under CCO care were immediately evaluated by a team member and then either directly cared for by the CCO team or by ward staff under their instruction. Stable patients were discharged from CCO after 72 hours. Patients who remained acutely ill and hemodynamically unstable or whose conditions caused concern were transferred to the intensive care unit.

Before randomization to the intervention arm (unexposed) wards, usual care continued. Ward nurses cared for acutely ill patients under the supervision of ward physicians. Physicians could request transfer to intensive care, but this was largely based on their individual judgment, rather than using scoring systems or formal referral criteria.

2.5. Outcomes

Primary outcomes were inhospital mortality and number of patients undergoing cardiopulmonary resuscitation (both expressed per patient). Secondary outcomes were length of stay and intensive care unit admission.

2.6. Data collection procedures

Data collection procedures were developed specifically for this evaluation. An independent data team was notified daily of new admissions to the study wards and on the same day reviewed patient records to collect information on patients’ age, sex, reason for admission (medical, scheduled or unscheduled surgery, or ward transfer) and data required for the Simplified Acute Physiology Score (SAPS) II [16]. No additional investigations were undertaken; any missing SAPS II data items were assumed to be normal.

Mortality and length of stay data were obtained from the hospital electronic information systems. Data on cardiopulmonary resuscitation and admissions to the intensive care unit were obtained from nursing office and CCO team records by the CCO team in exposed wards and by the independent data team in unexposed wards. For these outcomes, data collection was, therefore, not blind to exposure status. Where there was uncertainty, outcome data were rechecked by reviewing patient records.

2.7. Sample size

The sample size for this study was for the most part fixed by its design. That is to say, we used an opportunity to make a randomized evaluation of an intervention which was going to be rolled-out. Our study size was, therefore, constrained by the duration that it would take to roll-out the intervention to all wards. However, as preliminary power calculations suggested that this amount of data might only be able to detect larger differences, we added the 12 weeks preperiod and 12 weeks postperiod worth of data (calculations showed that any additional data had no material impact on power). Over the 72 weeks of the trial, 23,000 admissions to the wards were expected. We used Hussey and Hughes methods to calculate the minimum detectable effect based on the mortality rate (primary outcome) in those unexposed to the intervention and the magnitude of the intracluster correlation (ICC) of mortality rates [17]. With estimated inhospital mortality of 3.5%, ICC from 0.01 to 0.05, and an average cluster size of 1770, the study design would have 80% power (at 5% significance) to detect a decrease in mortality to 2.35 (a 35% relative risk reduction). This effect size is moderate to
large but smaller than the effect found in a study of similar design in the UK [18].

2.8. Randomization

The 13 wards were grouped into pairs (and 1 group of 3) with similar expected ward mortality rates. The 2 wards with the highest expected mortality rate were paired; the next 2 highest expected mortality wards were paired and so on. The 2 smallest wards had similar expected mortalities and were combined. For each pair, 1 ward was randomly allocated to initiate the intervention first in the first half of the study and the other second. The 6 pairs were then randomly allocated to their order in the sequence.

2.9. Allocation concealment and blinding

Randomization was carried out at a fixed point in time independent of the trial team, and the information on ward sequence was revealed 2 to 3 days before start of the transition period. Allocation concealment from individual patients was not important, as there was neither individual patient recruitment nor consent. The intervention was delivered without blinding.

2.10. Statistical methods

Admissions to wards were categorized as unexposed, transition phase, or exposed, and baseline characteristics (age, sex, type of admission, chronic diseases, and SAPS II score) were summarized by category.

We tested the null hypothesis of no difference in mortality rates before and after exposure using a mixed-effect logistic regression model. In addition to other patient or ward characteristics, we adjusted for clustering (ward), calendar time (since the intervention is sequentially rolled-out), and exposure to the intervention for each ward at each time point. We report the odds ratio (OR) as the intervention effect. The primary analysis was unadjusted except for clustering and time effects. A secondary analysis adjusted for prespecified patient covariates, age, sex, SAPS II score, and type of admission (elective or emergency).

Binary secondary outcomes were analyzed in a similar way. Because length of stay was markedly nonnormally distributed, we used a log-linear model and report exponentiated coefficients, which can be interpreted as the ratio of geometric means (or as a ratio of medians). These models were fitted using random-effects models in STATA, using the meglm function. As there were convergence difficulties using STATA, we used the Laplace approximation. We report the latent ICC, as is recommended in settings with binary outcomes and use the STATA function loneway [19].

All outcomes were considered significant at the 5% level, and we report both unadjusted and adjusted treatment effects, along with estimates of the ICC. The primary analysis was by intention to treat, with patients categorized based on the exposure status of the ward to which they were admitted. For the fully adjusted analysis, less than 2% of patients had incomplete data so missing data methods were not warranted.

2.11. Sensitivity analysis

Admissions during the transition phase are excluded from the main analysis. However, in a sensitivity analysis for the 2 primary outcomes, admissions during the 8-week transition phases were categorized as exposed to the intervention. Because patients transferring between wards might have been incompletely exposed to CCO, we also report outcomes subdivided by place of death.

2.12. Changes to methods after trial commencement

We had initially planned to fit the statistical models using population-averaged models, using generalized estimating equation methods, in STATA because random-effects models in cluster trials lack appropriate interpretation [20]. However, when model fitting, the generalized estimating equation failed to converge or took a very long time to run. We, therefore, used random-effects models. Where results did converge for both methods, overall conclusions were similar.

2.13. Qualitative evaluation

After implementation was completed, a nested qualitative study explored challenges to implementation and contextualized outcomes of the trial [21]. Between February and April 2012, 2 focus groups were conducted with nurses delivering the intervention and health professionals on the wards and followed up with individual interviews to clarify issues arising from focus groups. Audio recording of the groups was transcribed and translated from Farsi. Data collection ended when no new information emerged. Data were analyzed using an inductive content analysis approach [22]. AJ conducted the focus groups and interpreted the recordings; AL provided feedback on the developing thematic categories.

3. Results

3.1. Participant flow

Between July 2010 and December 2011, there were 22,919 admissions in the 13 wards: 1890 could not be included because the patient was discharged or transferred to another ward within 24 hours of admittance during the transition phase are excluded from the main analysis. However, in a sensitivity analysis for the 2 primary outcomes, admissions during the 8-week transition phases were categorized as exposed to the intervention. Because patients transferring between wards might have been incompletely exposed to CCO, we also report outcomes subdivided by place of death.

These models were fitted using random-effects models in STATA, using the meglm function. As there were convergence difficulties using STATA, we used the Laplace approximation. We report the latent ICC, as is recommended in settings with binary outcomes and use the STATA function loneway [19].

All outcomes were considered significant at the 5% level, and we report both unadjusted and adjusted treatment effects, along with estimates of the ICC. The primary analysis was by intention to treat, with patients categorized based on the exposure status of the ward to which they were admitted. For the fully adjusted analysis, less than 2% of patients had incomplete data so missing data methods were not warranted.

Admissions during the transition phase are excluded from the main analysis. However, in a sensitivity analysis for the 2 primary outcomes, admissions during the 8-week transition phases were categorized as exposed to the intervention. Because patients transferring between wards might have been incompletely exposed to CCO, we also report outcomes subdivided by place of death.

2.12. Changes to methods after trial commencement

We had initially planned to fit the statistical models using population-averaged models, using generalized estimating equation methods, in STATA because random-effects models in cluster trials lack appropriate interpretation [20]. However, when model fitting, the generalized estimating equation failed to converge or took a very long time to run. We, therefore, used random-effects models. Where results did converge for both methods, overall conclusions were similar.

2.13. Qualitative evaluation

After implementation was completed, a nested qualitative study explored challenges to implementation and contextualized outcomes of the trial [21]. Between February and April 2012, 2 focus groups were conducted with nurses delivering the intervention and health professionals on the wards and followed up with individual interviews to clarify issues arising from focus groups. Audio recording of the groups was transcribed and translated from Farsi. Data collection ended when no new information emerged. Data were analyzed using an inductive content analysis approach [22]. AJ conducted the focus groups and interpreted the recordings; AL provided feedback on the developing thematic categories.

3. Results

3.1. Participant flow

Between July 2010 and December 2011, there were 22,919 admissions in the 13 wards: 1890 could not be included because the patient was discharged or transferred to another ward within 24 hours of

<table>
<thead>
<tr>
<th>Month</th>
<th>Unexposed to intervention</th>
<th>Transition period (training)</th>
<th>Exposed to intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>7802</td>
<td>2347</td>
<td>10,880</td>
</tr>
<tr>
<td>Age (y)*</td>
<td>44 (20)</td>
<td>43 (19)</td>
<td>43 (19)</td>
</tr>
<tr>
<td>Male</td>
<td>3732 (46)</td>
<td>983 (42)</td>
<td>4266 (39)</td>
</tr>
<tr>
<td>SAPS II score</td>
<td>13.0 (3.8)</td>
<td>12.3 (9.3)</td>
<td>122 (9.6)</td>
</tr>
<tr>
<td>Type admission Scheduled surgery</td>
<td>2113 (27)</td>
<td>739 (31)</td>
<td>3849 (35)</td>
</tr>
<tr>
<td>Medical</td>
<td>3689 (47)</td>
<td>969 (41)</td>
<td>4124 (38)</td>
</tr>
<tr>
<td>Unscheduled surgery</td>
<td>1684 (22)</td>
<td>621 (26)</td>
<td>2855 (26)</td>
</tr>
<tr>
<td>Not known</td>
<td>316 (4)</td>
<td>18 (-1)</td>
<td>52 (-1)</td>
</tr>
<tr>
<td>Transferred ward</td>
<td>170 (2)</td>
<td>29 (1)</td>
<td>170 (2)</td>
</tr>
<tr>
<td>Chronic diseases AIDS</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (0.02)</td>
</tr>
<tr>
<td>Hematological</td>
<td>107 (1.37)</td>
<td>59 (2.51)</td>
<td>235 (2.16)</td>
</tr>
<tr>
<td>Metastatic cancer</td>
<td>35 (0.45)</td>
<td>3 (0.13)</td>
<td>36 (0.33)</td>
</tr>
</tbody>
</table>

Values are expressed as numbers and percentages.

* Mean (SD).
admission and before baseline data had been collected, leaving 21,029 patients contributing data to the trial (Table 1). In the unexposed phase, 7,802 were admitted; in the transition phase, 2,347; and in the exposed phase, 10,880. The CCO was implemented as intended under the randomization schedule in all study wards (Fig. 1). On 1,682 occasions the CCO took over the management of patients. On 46.2% of occasions, they were called by ward staff; on 53.2%, by members of the CCO team; and on 0.7%, after an emergency call. They also managed 879 patients after discharge from intensive care.

3.2. Baseline data

Age, sex, and SAPS II scores were similar in patients admitted during the unexposed and exposed periods. There were some differences in the reason for admission (Table 1). Small numbers of patients were transferred between the wards, 170 (2%), in both the unexposed and exposed periods.

3.3. Outcomes

There were 370 deaths (4.74%) among patients admitted to a ward during its unexposed period and 384 deaths (3.53%) among patients admitted to a ward during its exposed period: totally unadjusted OR, 0.73 (95% confidence interval [CI], 0.64-0.85) (Table 2). However, after adjusting for time effects, covariate effects and clustering this effect became more uncertain (adjusted OR, 1.02; 95% CI, 0.68-1.55). Mortality appears to decline over time in patients admitted to unexposed wards (Supplementary Fig. 2). The secondary outcome followed a similar pattern: the proportion of patients receiving cardiopulmonary resuscitation decreased from 4.86% to 3.61% (totally unadjusted OR, 0.73; 95% CI, 0.64-0.85). However, this result became more uncertain after adjusting for clustering and time effects (adjusted OR, 1.00; 95% CI, 0.69-1.48).

A similar proportion of patients was admitted to the intensive care unit in the 2 study periods (1.28% in the unexposed period and 1.23% in the exposed period), and the fully adjusted OR was 1.15 (95% CI, 0.64-2.00). Length of stay declined from a median of 6 days in the unexposed wards to a median of 4 days in the exposed wards. We do not report an adjusted effect for length of stay as the model was unstable and the results unreliable. All temporal trends in outcomes in unexposed wards were similar (Supplementary Fig. 2). Intracluster correlations were higher than anticipated, 0.013 (95% CI, 0.000-0.259) to 0.0979 (95% CI, 0.012-0.184).

3.4. Secondary analyses

Stratifying the results by the ward where the outcome event took place (ie, either on the admitted ward or on a transferred ward) showed similar findings to the primary analysis: mortality and cardiopulmonary resuscitation rates declined in crude analyses but became uncertain after adjustment (Supplementary Tables 3 and 4).
4.3. Limitations

The stepped wedge cluster randomized trial is a novel study design, and the design and analysis of these studies are still in development. Intraclass correlations were larger than expected, the effect of the intervention smaller, and there were large changes over time in the unexposed clusters. These all added to the wider CIs and the study's limited power to detect a small difference in mortality rates. The binary outcomes were modeled using the Laplace approximation which is not optimal, and there was such instability with the convergence of the models for length of stay that it was unreportable. A further limitation of our analysis is that there was a considerable amount of missing data in our analysis, and we would identify effects of the CCO team on ward patients not under their care. The nested qualitative evaluation may help explain why the intervention did not seem to change the outcomes.

4.4. Discussion

4.4.1. Summary of findings

This stepped wedge cluster randomized trial of CCO in a teaching hospital in a middle-income country did not show reductions in mortality and proportions of patients needing cardiopulmonary resuscitation or admitted to the intensive care unit or in length of stay. The qualitative study revealed resistance to the CCO service by ward staff because of perceptions of conflict with existing ward routines and increased workload. Difficulties in implementation may explain the lack of effect. Alternately, it may be that this study was underpowered because the intervention had a smaller than anticipated effect. We are uncertain whether this intervention is beneficial.

Qualitative evaluation suggests that both structural factors (existing staff workload and endorsement from hospital management) and attitudinal factors (willingness to change practice and understanding of end-of-life care) were not sufficiently in place for the intervention to succeed. There are particular ethical challenges in relation to end-of-life care in the context of following Islamic principles, and there are no guidelines for do-not-resuscitate orders in Iran [23]. However, do-not-resuscitate orders are found in other countries following Islamic principles [24].

4.2. Strengths

This was a well-designed study, avoiding many of the problems of simple before-and-after studies through the use of a stepped wedge cluster randomized trial with wards matched on predicted mortality rates. Simple before-and-after designs do not adequately deal with confounding [25]. Uncertain conclusions under a robust design are preferable to an erroneous conclusion of effectiveness [26,27]. As all admissions to the wards were included in the analysis, we would identify effects of the CCO team on ward patients not under their care. The nested qualitative evaluation may help explain why the intervention did not seem to change the outcomes.
covariate data, which meant that the fully adjusted estimate of the treatment effect was based on a relatively small subset of the observations. We did not use a multiple imputation because the methods of analysis for stepped-wedge studies are in their infancy and no methods currently exist for a multiple imputation of missing data from a stepped-wedge study.

Because of the need for bespoke data collection methods, the ascertainment of secondary outcomes was not blinded. We also lack quantitative data on important process measures such as numbers of patients reviewed by the CCO team. Ideally, a second researcher would have reviewed the interview and focus group transcripts.

This evaluation considered 4 outcomes: mortality, admission to the intensive care unit, the need for cardiopulmonary resuscitation, and length of stay. We did not consider resource costs or potential for harm, if diverting resources from other services affected other outcomes.

### 4.4. Comparison to existing literature

This is the only robust evaluation of CCO in a lower income or middle-income country. Our findings are consistent with the findings of evaluations of the effects of CCO teams in developed countries, which, when considering all the evidence, found no overall reduction in mortality but also observed substantial heterogeneity between studies [10]. The design of this study was modeled on a study of CCO implementation in the UK [18]. However, unlike the UK study, it found a large and statistically significant reduction in mortality. Some studies have found that ward staff did not call or delayed calling outreach teams, and this may explain low effectiveness [28]. In our study, resistance to the intervention by ward staff suggests that contextual factors may have impeded effectiveness in a similar way. It has been suggested that ongoing education of ward staff and review of the use of the CCO team over several years may improve the effectiveness of the CCO by changing organizational behavior; however, our study was not of sufficient duration for this to happen [29].

This pragmatic evaluation of implementation of CCO in a middle-income country teaching hospital did not find evidence of a reduction in mortality or other outcomes. Changes in health services in middle-income countries need robust evaluation. The stepped wedge study design is a feasible method of evaluation.

Supplementary data to this article can be found online at [http://dx.doi.org/10.1016/j.jcrc.2016.07.018](http://dx.doi.org/10.1016/j.jcrc.2016.07.018).

### Acknowledgments

Much gratitude is due to Shariati Hospital and Tehran University of Medical Sciences for their generous assistance in supporting this research. The researchers are very grateful to the Shariati hospital staff for their genuine cooperation and trust.

### References


[27] Moberly T. Service redesign should be tested as rigorously as new treatments, NHS chief says. BMJ 2014;348:g3744.
