Author's response to reviews

Title: Early warning scoring systems versus standard observations charts for wards in South Africa: a cluster randomised controlled trial

Authors:

Una Kyriacos (una.kyriacos@uct.ac.za)
Jennifer Jelsma (jennifer.jelsma@uct.ac.za)
Sue Jordan (s.e.jordan@swansea.ac.uk)

Version: 2 Date: 6 January 2015

Author's response to reviews: see over
Early warning scoring systems versus standard observations charts for wards in South Africa: a cluster randomised controlled trial

Thank you for your e-mail of 10 December 2014 and referee reports. We are grateful for the constructive suggestions of the reviewers and editor. These have been adopted. We agree that they strengthen the manuscript; however, responding to these important points has lengthened the paper. The changes are tabulated below.

I am pleased to say that Professor Mike James has returned to work following a critical illness. He contributed to the study design and is now able to contribute to the final draft of the paper. Please could you let us know if it is acceptable to include an additional author?

Please contact me if clarification is needed.

Thank you for your help.

Una Kyriacos
### Responses to Reviewers

<table>
<thead>
<tr>
<th>Comments</th>
<th>Response</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Major Compulsory Revisions</strong></td>
<td>Both reviewers have recommended revision of outcomes. Thank you. We agree.</td>
<td></td>
</tr>
<tr>
<td>Reviewer 1</td>
<td>The primary outcome is now: Nurses’ documented responses to abnormal vital signs measures: rechecking vital signs or calling for assistance from a more senior nurse or medical doctor for patients who met pre-selected criteria for intervention.</td>
<td>9</td>
</tr>
<tr>
<td>Three primary outcome is too much. I would recommend to revise this and exclude the nurse knowledge as a primary outcome. To me, this is more of a quality control measure to insure that nurses randomised to the intervention received the intervention. While I agree that it is important to insure that nurses received a proper training, I do not think this should the primary (most important) outcome of the study.</td>
<td>Tables have been renumbered and reordered to first present record review data, followed by knowledge test score results.</td>
<td></td>
</tr>
<tr>
<td>Reviewer 2</td>
<td>The rest are now secondary outcomes. We have re-ordered tables accordingly.</td>
<td>9</td>
</tr>
<tr>
<td>[1] Better specification of vital sign outcomes: currently stated as “the number of patients with physiological variables recorded on the MEWS chart in intervention wards and on existing vital signs charts in control wards”.</td>
<td>[a] the time is now specified as: in the first 8 postoperative hours;</td>
<td>9</td>
</tr>
<tr>
<td>This omits [a] specification of time frame,</td>
<td>[b] the actual variables of interest are now listed: (respiratory and heart rate, oxygen saturation, systolic blood pressure, temperature, level of consciousness and urine output);</td>
<td>9</td>
</tr>
<tr>
<td>[b] the actual variables of interest,</td>
<td>[c] reference to abnormal vital signs measures is now captured in the primary outcome.</td>
<td>9</td>
</tr>
<tr>
<td>Moreover the Number of times each vital sign was recorded in the first 8 hours may be of interest to readers wanting to understand practice.</td>
<td>A further outcome has been added: The number of times each vital sign was recorded in the first 8 postoperative hours in both trial arms.</td>
<td>9</td>
</tr>
<tr>
<td>Was one a primary outcome.</td>
<td>(This is covered in the response to reviewer 1, above)</td>
<td>9</td>
</tr>
<tr>
<td>[2] Better specification of clinical response outcomes: currently stated as “nurses’ responses to high and low threshold MEWS recordings in patients’ notes.” I think it would be helpful to clarify what responses were sought, to present these responses, and to explain how/if medical responses were included. Also the activation of ICU or similar team would be of interest.</td>
<td>Expected responses are now outlined in the reworded primary outcome, excluding activation of ICU or outreach teams. (This is covered in the response to reviewer 1, above)</td>
<td>9</td>
</tr>
<tr>
<td><strong>Under Study sites</strong></td>
<td>Under Study sites we have added: The traditional ‘cardiac arrest team’ comprising ICU nurses and doctors, had been replaced by ward response teams more than two decades previously. There was no hospital-wide emergency response system for predefined thresholds for deterioration in physiological variables and no early warning scoring system in place on any general wards.</td>
<td>5</td>
</tr>
</tbody>
</table>
Our Mission is to be an outstanding teaching and research university, educating for life and addressing the challenges facing our society.

| [3] Interpretation of MEWS scores / sub-scores: the authors state 'needed to be assessed'. I believe the statement 'met pre-selected criteria for intervention' rather than 'need'. Indeed it is not apparent from the data that the patients with one or more vital signs in the 'requires intervention range' missed out on important care. | This is now captured:  
- In results of the abstract: Record review offered no statistically significant evidence that introduction of the MEWS was associated with summoning assistance for patients who met pre-selected criteria for intervention (50/57 received no assistance, vs 55/57, OR 0.26, 0.05-1.31) despite improvement in nurses' knowledge in intervention wards.  
- in the primary outcome: Nurses’ documented responses to abnormal vital signs measures: rechecking vital signs or calling for assistance from a more senior nurse or medical doctor for patients who met pre-selected criteria for intervention.  
- Responses to MEWS triggers that should have been reported for patients who met pre-selected criteria for intervention are summarised in Table 3. | 2 |

| [4] Table 4 provides rich data that for me is the core of this study. I suggest it can be simplified by [a] excluding duplication of the was and was not recorded;  
[b] use of more columns for the should have triggered, 'did trigger' ;  
[c] removal of the non-p-value statistical items – these can be determined by repeating the calculations if the reader desires;  
[d] clarify what 'trigger' means in the context of this research – ie. documentation of X,Y,Z etc. | Renumbered as Table 6.  
[a] the data row 'was not recorded' has been deleted.  
[b] Done  
[c] We are uncertain as to the meaning of this. We are happy to remove P values where P =>0.05, but are not sure this is what is intended so have retained these in tables. Readers can easily calculate ORs, but it is less easy to calculate 95% confidence intervals without recalling certain formulae. (In our experience, many nurses do not easily recall these formulae). The confidence intervals are essential for interpretation of the data, most particularly to allow readers to interpret any uncertainties. We agree with Professor Altman (1991 p.175) that the presentation of both P values and confidence intervals is desirable, and if only one can be presented, it should be confidence intervals.  
[d] Table 4: We have added a Note on table: † MEWS trigger of 1 = recheck measurement after ½ hour and report if no improvement; MEWS trigger of 2 = recheck measurement after 5 minutes and report immediately if no improvement; MEWS trigger of 3 = critical, report urgently. | 30 |
```
Table 5 seems closely related, and may be able to be combined with table 4.

Table 5 is renumbered as Table 3. Table 4 is renumbered as Table 6. We have not merged these 2 tables: Table 6 is the per protocol analysis of only 36 records where the MEWS was used in the intervention wards (the existing chart was used for i.e. usual care given to) 21 patients in the intervention ward) and all the records (N=57) in the control wards whereas data analysis in Table 3 is for intention to treat i.e. all records.

[5] Interpretation: The results of the study are intriguing, in that there were differences in documentation, but similar rates of ‘abnormal’ vital signs (by MEWS criteria) documented. This highlights the question of the clinical utility of the MEWS system to identify patients at risk. The data suggest transgression of threshold was common and documented responses were infrequent. The absence of routinely catastrophic clinical consequences seems noteworthy. Expanding the discussion about these findings would strengthen the manuscript— as would including the rationale for changing to another system.

[5] Thank you for these suggestions that strengthen the manuscript. We have made certain changes and expanded the discussion.

This raises questions regarding the clinical effectiveness of the MEWS chart deployed to identify patients at risk and the user friendliness of the chart. To address the latter possibility, a revised version of the Cape Town ward MEWS observations chart (Supplementary Figure 2: Revised Cape Town MEWS chart) is therefore presented.

[6] Discussion: Discourse on delegation of vital signs. The observation that no scores were calculated is concerning; and should be further discussed. What was the plan for follow-up and audit after initial implementation. This (in retrospect) seems a limitation of planning in this 5 month study.

[6] We have expanded the discussion. We have added a further limitation that researchers had no authority to plan for implementation of the EWS system at the conclusion of the study. We have added a reference for the absence of a total MEWS.

[7] Implementation Effectiveness: this is an important, but under developed part of the paper. It appears the intervention was not used. 63% of the intervention patients did not use the new charts (why?).

[7] This is a misunderstanding: the MEWS chart was used in 63% of patients in intervention wards. Some logistical reasons for the 37% not having the chart have been added to the text.

The Cape Town ward MEWS has a wider range of cut points for respiratory rate, oxygen saturation, heart rate, systolic BP and temperature than other published MEWS. This might have made interpretation of parameter readings more difficult and may have accounted for the low response rate to the large number of triggers, despite training.

Anecdotal reports at the conclusion of the trial of nurses from control wards, untrained in the MEWS, voluntarily working overtime shifts in intervention wards thereby unwittingly delivering some aspects of the intervention would have diluted the impact of the intervention [reference given]: this may account for some, but not all, of the 37% non-completion of MEWS scores and the improvement in nurses’ knowledge in control wards.
```
| Scores were calculated by 'no-one', and documentation of responses for patients above threshold were limited. This is a fundamental finding, that warrants much greater component of the discussion. | No scores were totalled. Complex systems such as EWS are prone to human calculation errors, particularly use of the traditional ‘pen and paper’ method [ref given] compared with technology [ref given]. Not all patients in both trial arms had recordings of all 7 vital signs. In our previous work [reference given] in the same six wards, results showed a similar pattern of recording of vital signs and that no patients’ records contained recordings for all seven parameters displayed on the MEWS. The absence of a total MEWS or one or more (of 6) vital signs’ recordings in 47% (n=509) of patients in general surgical and medical wards in an observational data analysis pre and post intervention study in Belgium [reference given] resonates with our findings. There were differences in recordings between trial arms but similar rates of ‘abnormal’ vital signs (by MEWS criteria). | 14 |

| [8] Appendix: Inclusion of the actual questions in the educational assessment would be of interest and may help the reader to better interpret the clinical knowledge both before and after education. For me 60% seems a low post-test score and a 20% increment is also modest. Understanding the test would help here. | [8] Appendix 1 has been added for the editor’s discretion. See additional text. |  |

| [9] Randomization: please can the authors better separate randomization of the intervention from randomly selected patients for record review. The process may be similar, but should be reported for intervention and outcome assessment separately. | [9] A separate heading has been added. Changes made as follows: Randomisation and sequence generation of wards: Six wards were randomised to either the intervention or control wards by drawing of sealed lots [ref]. | 7 |

| [10] Analyses: these presently do not take account for clustering. | Yes, acknowledged as a limitation of the study. Low numbers preclude adjusted analyses. | 15 |

**Minor Essential Revisions**

**Reviewer 1:**

1. Abstract
There should be a designation of the primary outcome and information about statistical analysis in the method section of the abstract.

Done
Methods:
2. In the sampling section, the authors should mention whether nurses worked only on a single ward or if it was possible to have a nurse working on different ward. This later situation could contaminate the intervention.

| Methods: 2. In the sampling section, the authors should mention whether nurses worked only on a single ward or if it was possible to have a nurse working on different ward. This later situation could contaminate the intervention. | Thank you. We have now reported this more explicitly as: A permanent nurse was unlikely to be assigned to both an intervention and control ward therefore allocation of all nurses in respective clusters to the intervention or control arm rather than individual nurses limited the threat of contamination. Unlike agency staff, permanent nurses were assigned to single wards, reducing the threat of contamination. Addressed in 7 above: in the discussion we have added: Anecdotal reports at the conclusion of the trial of nurses from control wards, untrained in the MEWS, voluntarily working overtime shifts in intervention wards thereby unwittingly delivering some aspects of the intervention would have diluted the impact of the intervention. | 5 |

3. The authors should justify their exclusion criteria. For example, why exclude patients who died or those who were transferred to ICU in the first 8 hours? They may have die because of improper management related to the absence of MEWS.

| 3. The authors should justify their exclusion criteria. For example, why exclude patients who died or those who were transferred to ICU in the first 8 hours? They may have die because of improper management related to the absence of MEWS. | We have expanded this section to read: Exclusion criteria for patient records: 1) Absent, incomplete or unavailable records (for example due to medico-legal review). Incomplete records were defined as not including observations charts and patient progress notes. If either was absent, case-note were discarded, and replaced by an alternative random number. 2) Case-notes of patients designated ‘not for resuscitation’ as this might have affected nurses’ responses to changes in vital signs. 3) Case notes of patients transferred to high dependency care from the operating room or from the ward within 8 hours of their operation, including patients who had died there, because the study focus was ward-based vital sign monitoring. | 6 |

4. The sample size calculation is suboptimal. Previous studies reported respiratory rate monitoring of 40% while it was 2% in preparatory work. The sample size was calculated using the 2% but it would be much larger using 40%. Also, as mentioned, the effect of clustering was not taking into account in the sample size calculation. This effect would also increase the sample size.

| 4. The sample size calculation is suboptimal. Previous studies reported respiratory rate monitoring of 40% while it was 2% in preparatory work. The sample size was calculated using the 2% but it would be much larger using 40%. Also, as mentioned, the effect of clustering was not taking into account in the sample size calculation. This effect would also increase the sample size. | Thank you. This limitation had been acknowledged. We took a locally derived figure (2%), rather than the 40% observed in the UK. As in all trials, we were unable to take account of clustering by practitioner. | 6 |
5. What was the delay between the pre and post knowledge testing?

| Staggered pre-intervention knowledge testing was conducted between 12 March and 23 March for nurses in both trial arms. The MEWS training programme was conducted between 12 March and 22 April. The MEWS chart was introduced on 1 May and removed at midnight on 31 July. Post-intervention knowledge testing was conducted between 3 August and 13 August. We should like the editor’s advice as to whether this detail should be included in the text. We have added it as a footnote to the questionnaire (Appendix 1). |

6. The author showed a difference in vital signs recording post intervention but it is possible that there was a difference before. The author could look at a sample of charts before intervention in both group to demonstrate that there was no difference before. They did it to compare nurses’ knowledge, they could do it for vital signs.

| Added: Our previous work [reference given] in the same six wards indicated a similar uniform absence of responses to abnormal vital signs. |

7. The fact that the nurses failed to report deranged physical parameter is worrisome. However, it looks like 87.7% (intervention) and 96.5% (control) of the patients had deranged parameters. This is very high and makes may doubt about the specificity of the MEWS to identify “at risk” patients. This should be commented in the discussion

| We agree. We have added: However, in our previous work [reference given] in the same research wards, there were no reported interventions for 10/11 (90.9%) patients who died and had abnormal physiology, and for 38 (86.4%) controls with abnormal physiology. Specificity of the Cape Town MEWS had been determined at between 77.3% and 81.4% [reference given]. Strategies for implementing and monitoring patient safety policies in a teaching hospital ought to consider the complexity of such an organisation. Patients admitted to this level of care have complex conditions requiring treatment not available at secondary level hospitals that increase their risk of an AE regardless of their comorbidities. |

Discussion

7. Benefits of intervention. I do not agree with the authors claiming the importance of the training program. This study showed an improvement of the reporting of vital signs but no impact on reports. My understanding was that the training program was to improve nurses’ knowledge of physiology so they can react faster to abnormal vital signs. My understanding of the results is that the MEWS improves vital signs measurement but has no impact on reporting. This latter may be related to the training.

| We agree and have added the following: The improved knowledge scores disappointingly only reached 61% and may have been impacted by language difficulties and the inclusion of non-professional nurses who have a non-academic qualification. The MEWS chart resulted in significantly more patients in the intervention wards than control wards having recordings of respiratory rate, the best discriminator for clinical outcomes [reference]. Not all existing observation charts [reference] included respiratory rate monitoring. This trial showed that if included, respiratory rate is more likely to be monitored. Because respiratory rate monitoring is the best discriminator of serious adverse events (SAEs), it has replaced recording of oxygen saturation on |
some charts [ref]. Nevertheless, the improved scores suggest that a hospital-wide training programme for early recognition and management of patients with impending critical illness is needed to complement in-service education programmes for late rescue techniques such as cardiopulmonary resuscitation (CPR). Low pre-intervention bioscience knowledge scores in both trial arms confirm previous research [ref] that suggests that this underpinning knowledge may be suboptimal.

Improved recording did not translate into an improvement in responses to deranged physiology. The Cape Town ward MEWS has a wider range of cut points for respiratory rate, oxygen saturation, heart rate, systolic BP and temperature than other published MEWS. This might have made interpretation of parameter readings more difficult and may have accounted for the low response rate to the large number of triggers, despite training. This raises questions regarding the clinical effectiveness of the MEWS chart deployed to identify patients at risk and the user friendliness of the chart. To address the latter possibility, a revised version of the Cape Town ward MEWS observations chart (Supplementary Figure 2: Revised Cape Town MEWS chart) is therefore presented. The UK National Early Warning Score (NEWS) [ref] was not available when the study was undertaken. Since completion of this trial, we have substituted the values from the NEWS into the Cape Town ward MEWS with permission (Royal College of Physicians) as the NEWS values retain the range of trigger cut points (heart rate 111-129; systolic BP 81-100) we found were associated with mortality [ref]. User friendliness of the chart might be improved by reducing the number of cut points for physiological parameters that were not associated with mortality. To assist those responding to patients’ physiological deterioration, the UK NEWS algorithm for a total MEWS has been included on the chart in addition to guidelines for reporting a single parameter trigger.

Tables and figure
8. Tables 3 and 4 could be simplified by excluding the following row: RR not recorded, HR not recorded, Oxygen sat not recorded, etc…

Done – now Tables 5 and 6 respectively. 29 30
**Reviewer 2:**

**Minor points / Discretionary Revisions**

[a] I am not convinced by the argument that random selection of patients was needed to reduce Nurse Hawthorne effect – but am convinced it was a pragmatic approach to selecting patients for outcome assessments.

[b] Analysis of deaths – could be included;

[c] consider describing what the recorded interventions were.

[a] Reference to the Hawthorne effect has been deleted.

[b] Analysis of the records of the 3 patients who died is included in the sample and not presented separately to preclude patient identification. The 3 deaths occurred in the intervention wards as shown in Figure 1.

[c] Consider describing what the recorded interventions were.

---

**Discretionary Revisions**

**Reviewer 1:**

1. Results

An improvement from 10/23 to 14/23 seems small following training in the intervention group. Was there any explanation for such a low score following training?

Some background is provided. The improvement in scores between pre- and post-intervention knowledge tests of nurses in the intervention arm to 61% did not translate into improved documented reporting of patients who met predetermined criteria for intervention. Nurses failed to report deranged physiological parameters for 87.7% (intervention) and 96.5% (control) patients. This is deeply concerning and raises questions about the specificity of the MEWS to identify “at risk” patients. However, in our previous work [1] in the same research wards, there were no reported interventions for 10/11 (90.9%) patients who died and had abnormal physiology, and for 38 (86.4%) controls with abnormal physiology. Specificity of the MEWS had been determined at between 77.3% and 81.4%. Patients’ progress notes revealed few entries of nurses’ responses to signs of physiological deterioration and this was interpreted as inappropriate responses to MEWS that should have triggered a call for assistance. Other studies also report nurses’ failure to communicate concerns and inappropriate responses when patients showed signs of physiological deterioration [3, 8, 51]. Other factors have to be considered. The MEWS does not replace clinical judgement in detecting deteriorating patients. Although there were differences in recordings between trial arms but similar rates of “abnormal” vital signs (by MEWS...
criteria), the absence of routinely catastrophic clinical consequences in both trial arms is noteworthy but not unusual [29]. There are too many confounding factors in a clinical setting to attribute these outcomes to inadequate recording and reporting. Strategies for implementing and monitoring patient safety initiatives in large tertiary hospitals need to accommodate the complexity of such organisations. Patients admitted to this level of care have complex conditions requiring treatment not available at secondary level hospitals, increasing their risk of adverse events.

**Editorial requests:**
1. Please include the email addresses of all authors on the title page.  
   **Done**  
2. Please include the date of registration with the trial registration number at the end of the Abstract.  
   **Done**  
3. Please rename the Introduction, 'Background'.  
   **Done**  
4. Please move the funding information to the Acknowledgements section.  
   **Done**  
5. Please include a figure title and legend section after the reference list.  
   **Done**  
6. Please include an additional file title and legend section after the figure legend section.  
   **Done**

**Reference**