Reviewer's report

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

Version: 1  Date: 12 October 2014

Reviewer: Christopher Parshuram

Reviewer's report:

Thank you for the opportunity to review the manuscript of Kyriacos et al. Described is a cluster randomized controlled trial design evaluating education and process outcomes associated with implementation of a complex intervention. Patients were 14 years or older, had received general anaesthetic, had no treatment limitations, and were treated in one of the six evaluated inpatient wards in a single center (Cape Town).

The intervention was MEWS (modified early warning score) with MEWS education and new observation chart. An education program was developed, staff education completed, and the MEWS chart was then used in the 3 allocated wards.

Outcomes were the education test scores (n=50 nurses), the number of patients with vital sign measurements in the first 8 post-operative hours (n=114 patients), and documented responses for observation that were ‘above threshold’ (MEWS score of 3).

Key findings were increased test scores after education, (with similar baseline scores), increased RR recording in intervention wards but similar HR and BP measurement rates; similar rates of abnormal vital sign documentation; and low rates of response to ‘above threshold’ patients – in both control and intervention arms. The MEWS score were calculated by ‘no-one’ and charts were routinely not used (~40%). Mortality was reported (appeared to be higher in intervention arm) but was not formally evaluated.

The authors report changing to another severity of illness score after the trial.

Major Compulsory Revisions:

[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”. This omits [a] specification of time frame, [b] the actual variables of interest, [c] the ‘normality’ abnormality of measures. 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. Was one a primary outcome.

[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.

[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.

[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 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. Table 5 seems closely related, and may be able to be combined with table 4.

[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.

[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.

[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 ?), 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.

[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.

[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

[10] Analyses: these presently do not take account for clustering.

Minor points / Discretionary Revisions
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.

Analysis of deaths – could be included;

consider describing what the recorded interventions were.

Thank you for the opportunity to review this work. I believe that with some revision it could provide an important addition to the literature on implementation fidelity, processes of clinical implementation and the clinical utility of implemented scores.

Sincerely

C Parshuram

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I hold shares in Bedside Clinical Systems, a clinical decision support company that has products (presently for children) that include an FDA approved early warning system. This company is unlikely to benefit financially from the publication - or non-publication of this manuscript. If the company decides to expand to adult EWS then this paper may be of relevance, however the recommendations I have made for this manuscript are more about implementation and the validity of the system used, than recommending a product.

Accordingly I believe I have provided an unbiased review, and unbiased recommendations.

Sincerely

Chris Parshuram.