Title	The effect of adult Early Warning Systems education on nurses' knowledge, confidence and clinical performance: A systematic review
Authors	Saab, Mohamad M.;McCarthy, Bridie;Andrews, Tom;Savage, Eileen;Drummond, Frances J.;Walshe, Nuala;Forde, Mary;Breen, Dorothy;Henn, Patrick;Drennan, Jonathan;Hegarty, Josephine
Publication date	2017-05-22
Original Citation	Saab, M. M., McCarthy, B., Andrews, T., Savage, E., Drummond, F. J., Walshe, N., Forde, M., Breen, D., Henn, P., Drennan, J. and Hegarty, J. (2017) 'The effect of adult Early Warning Systems education on nurses' knowledge, confidence and clinical performance: A systematic review', Journal of Advanced Nursing, 73(11), pp. 2506-2521. doi: 10.1111/jan.13322
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1111/jan.13322
Rights	© 2017 John Wiley & Sons Ltd. This is the peer reviewed version of the following article: Saab MM, et al. The effect of adult Early Warning Systems education on nurses' knowledge, confidence and clinical performance: A systematic review. J Adv Nurs. 2017;73:2506–2521, which has been published in final form at https://doi.org/10.1111/jan.13322. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.
Download date	2024-04-20 07:42:41
Item downloaded from	https://hdl.handle.net/10468/5370

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

MR MOHAMAD M SAAB (Orcid ID : 0000-0002-7277-6268)

Article type : Review

The Effect of Adult Early Warning Systems Education on Nurses' Knowledge, Confidence, and Clinical Performance: A Systematic Review

Running head: Early Warning Systems Education for Nurses: A Systematic Review

Mohamad M. SAAB, Lecturer, PhD(C), MSc, RN [a]

Bridie McCARTHY, Lecturer, PhD, RNT, RGN [a]

Tom ANDREWS, Lecturer, PhD, RNT, RGN [a]

Eileen SAVAGE, Professor, PhD, BNS, RGN [a]

Frances J. DRUMMOND, Primary Investigator, PhD [a,b]

Nuala WALSHE, Clinical Skills Simulation Centre Manager, PhD(C), RGN, RNT [a]

Mary FORDE, Nurse Practice Development Co-ordinator, DN(C), RGN, RM [c]

Dorothy BREEN, Director of Education, FFARCSI, FCICM, MD [d,e]

Patrick HENN, Lecturer, MB, MSc, MA [e,f]

Jonathan DRENNAN, Professor, PhD, MEd, RN [a]

Josephine HEGARTY, Professor, PhD, RNT, RGN [a,*]

a Catherine McAuley School of Nursing and Midwifery, University College Cork, Cork, Ireland

[b] Department of Epidemiology and Public Health, University College Cork, Cork, Ireland

[c] Nurse Practice Development Unit, Bon Secours Hospital, Cork, Ireland

[d] Intensive Care Department, Cork University Hospital, Cork, Ireland

[e] Application of Science to Simulation-based Education and Research on Training (ASSERT), University College Cork, Cork, Ireland

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jan.13322

[f] School of Medicine, University College Cork, Cork, Ireland

[*] Corresponding author at: School of Nursing and Midwifery, University College Cork, Brookfield Health Sciences Complex, College Road, Cork, Ireland.

Tel.: +353 21 490 1462

E-mail address: J.hegarty@ucc.ie

Funding Information: This review received funding from the National Clinical Effectiveness Committee, Department of Health, Ireland.

Conflict of Interest: No conflict of interest to declare

ABSTRACT

Aims: This review aims to determine the effect of adult Early Warning Systems education on nurses' knowledge, confidence and clinical performance.

Background: Early Warning Systems support timely identification of clinical deterioration and prevention of avoidable deaths. Several educational programmes have been designed to help nurses recognise and manage deteriorating patients. Little is known as to the effectiveness of these programmes.

Design: Systematic review

Data sources: Academic Search Complete, CINAHL, MEDLINE, PsycINFO, PsycARTICLES, Psychology and Behavioral Science Collection, SocINDEX and the UK & Ireland Reference Centre, EMBASE, the Turning Research Into Practice database, the Cochrane Central Register of Controlled Trials (CENTRAL) and Grey Literature sources were searched between October - November 2015.

Review methods: This is a quantitative systematic review using Cochrane methods. Studies published between January 2011 - November 2015 in English were sought. The risk of bias, level of evidence and the quality of evidence per outcome were assessed.

Results: Eleven articles with ten studies were included. Nine studies addressed clinical performance, four addressed knowledge and two addressed confidence. Knowledge, vital signs recording and Early Warning Score calculation were improved in the short-term. Two interventions had no effect on nurses' response to clinical deterioration and use of communication tools.

Conclusion: This review highlights the importance of measuring outcomes using standardised tools and valid and reliable instruments. Using longitudinal designs, researchers are encouraged to investigate the effect of Early Warning Systems educational programmes. These can include: interactive e-learning, on-site interdisciplinary Early Warning Scoring systems training sessions and simulated scenarios.

Keywords: Resuscitation; knowledge; confidence; clinical performance; clinical deterioration; education; nursing; systematic review; literature review.

SUMMARY STATEMENT

Why is this research or review needed?

- Acutely ill patients are at risk of developing adverse events leading to clinical deterioration, transfer to intensive care units and avoidable death.
- Well-established programmes exist to educate nurses about the use of Early Warning Systems in the recognition of clinical deterioration.

- There has been little attempt to systematically review recent evidence on the effectiveness of adult Early Warning Systems education in enhancing nurses' knowledge, confidence and clinical performance.

What are the key findings?

- There is a wide variation in the programmes used to educate nurses about Early Warning Systems.
- Results from this review indicate that Early Warning Systems education is effective in enhancing nurses' knowledge and confidence and clinical performance in the short-term.
- Several non-validated, researcher-designed tools were used to measure outcomes.

How should the findings be used to influence policy/practice/research/education?

- This review provides researchers with valuable information to select and/or develop outcome-based training programmes aimed at enhancing knowledge, confidence and clinical performance in relation to Early Warning Systems.
- Future studies must be interdisciplinary, delivered frequently and measured longitudinally.

INTRODUCTION

Acutely ill patients with complex health needs are increasingly being cared for on general wards and hence are at risk of clinical deterioration leading to adverse events such as cardiac arrest, transfer to intensive care units (ICU) and unexpected and avoidable death (Taenzer *et al.* 2011). In most cases, these adverse events are preceded by clinical signs of deterioration (Harrison *et al.* 2005, Jamieson *et al.* 2008, Fagan *et al.* 2012).

Timely detection and appropriate interventions are critical to providing safe and effective care to a clinically deteriorating patient (Alam *et al.* 2015). This involves identifying and classifying the severity of illness, providing prompts and structured communication tools to escalate care and following a definite escalation plan (National Clinical Effectiveness Committee [NCEC] 2013).

Failure to detect early signs of deterioration in the acutely ill patient is considered a major cause of avoidable morbidity and mortality (Alam *et al.* 2015). Attempts to achieve earlier identification of the clinically deteriorating patient led to the introduction of Early Warning Systems (EWS) in acute care settings (NCEC 2013).

Background

EWS, also known as track and trigger systems, are designed to facilitate early detection and communication of clinical deterioration by categorising the severity of illness and prompting timely review by the appropriately trained personnel at specific trigger points (Mitchell *et al.* 2010). EWS are based on an aggregate scoring system, where a score is allocated to key physiological parameters, including respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate and level of consciousness (Urban *et al.* 2015). The score allocated to each of the parameters is considered as a trigger point. For example, using the Irish National Early Warning Scoring (NEWS) system, a score of 3 on any of the aforementioned parameters serves as a trigger point, which requires healthcare

professionals to escalate care (NCEC 2013). Other examples of widely used track and trigger systems include the Modified Early Warning Scoring system (MEWS) (Urban *et al.* 2015) and the Bispebjerg Early Warning Scoring system (BEWS) (Christensen *et al.* 2011).

The introduction of EWS to adult general wards is complex (Robb & Seddon 2010). In addition, the effectiveness of EWS initiatives depends on the availability of adequate resources, leadership and healthcare professionals' knowledge and ability to recognise clinical deterioration. Several barriers to timely recognition and response to clinical deterioration exist. These include: lack of understanding of physiological deterioration and triggering criteria (De meester *et al.* 2013); failure to undertake complete and reliable vital sign measurement; incorrect calculation of aggregate scores (Ludikhuize *et al.* 2012); ineffective communication (Rabol *et al.* 2011); poor clinical reasoning skills (Levett-Jone *et al.* 2010); and inter-professional hierarchical factors such as the power relationships that exist between nurses and physicians (Shearer *et al.* 2012).

Several educational programmes have been designed to help nurses recognise and manage deteriorating patients, including Acute Life Threatening Events Recognition and Treatment (ALERT); Multi-professional Full-scale Simulation (MFS); COMPASS (Mitchell *et al.* 2010); and Acute Illness Management (AIM) (Liaw *et al.* 2011). Although there is a growth of educational programmes on various EWS, there has been little attempt to date to systematically review the recent evidence on their effect on nurses' knowledge, confidence and clinical performance.

THE REVIEW

Aims

The aim of this systematic review was to determine the effect of adult EWS education on nurses' knowledge, confidence and clinical performance. The search was conducted based on three pre-specified questions developed using the PICO (population, intervention, comparison, outcome) framework outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green 2011). These include:

(i) What is the effect of EWS educational programmes on nurses' level of knowledge, compared with baseline and/or control conditions?

(ii) What is the effect of EWS educational programmes on nurses' level of confidence compared with baseline and/or control conditions?

(iii) What is the effect of EWS educational programmes on nurses' clinical performance in terms of vital sign recording, EWS calculation and/or escalation of care compared with baseline and/or control conditions?

Design

A systematic review of educational interventions aimed at enhancing nurses' knowledge, confidence and clinical performance regarding EWS was undertaken. This review was guided by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green 2011) and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (Moher *et al.* 2009).

Search methods

Eligibility criteria

The PICOS (S for study design) framework was used to determine the study eligibility criteria (Moher *et al.* 2009). Randomised controlled trials (RCTs), non-RCTs and pre- and post-test studies considered for inclusion met the following criteria: (i) involved adult patients (i.e. over 18 years of age); (ii) comprised programmes relating to the education and/or training of nurses about the use of EWS/track and trigger systems; (iii) compared the effect of educational programmes to baseline and/or control conditions; (iv) addressed nurses' knowledge, confidence, an clinical performance in terms of vital sign recording, early warning score calculation and/or response to clinical deterioration; (v) published between January 2011 and November 2015; and (vi) published in English.

Studies with paediatric patients (i.e. aged less than 18 years) and/or pregnant patients were excluded as the scoring systems used in these patient populations are different. Opinion papers, policy reports, abstract-only articles, economic papers relating to budget impact analysis of EWS, studies evaluating the implementation of EWS and papers on the clinical effectiveness and validation of EWS were not deemed eligible for inclusion.

Search strategy

A systematic search of several electronic databases and the Grey Literature was conducted between October and November 2015. The databases searched were: Academic Search Complete; CINAHL; MEDLINE; PsycINFO; PsycARTICLES; Psychology and Behavioral Science Collection; SocINDEX; the UK & Ireland Reference Centre; EMBASE;

the Turning Research Into Practice (TRIP) database; and the Cochrane Central Register of Controlled Trials (CENTRAL). The Grey Literature searched included several guideline websites and repositories namely: Open Grey (2015); New York Academy of Medicine (2015); OpenDoar (2015); National Institutes of Health (2015); Health Service Executive (2015); Health Information and Quality Authority (2015); Health Research Board (2015); Lenus (2015); World Health Organization (2015); National Institute for Health and Care Excellence (2015); Department of Health (2015); National Health Service England (2015); Public Health Agency of Canada (2015); Google Scholar (2015); and Google (2015). The search was limited to studies published in English between January 1st, 2011 and November 30th, 2015. The reference lists of included studies were also searched for potentially eligible studies. The reason for limiting the search to five years was to capture the latest evidence, especially that new EWS are emerging and guidelines regarding staff education on the use of these systems are being continuously updated.

The PICOS framework was used to select and combine the search terms in a way that addressed the aim of this systematic review (Higgins & Green 2011). Keywords were searched on title and abstract and combined using Boolean operators 'AND' and 'OR' as well as Medical Subject Headings (MeSH). The full search strategy and eligibility criteria are presented in Table S1 in the online version of the article.

Risk of bias and level of evidence assessment

The risk of bias for each study, the quality of evidence for each review outcome and the level of evidence for each study were assessed. The risk of bias for RCTs and non-RCTs was assessed using the nine criteria of the Effective Practice and Organisation of Care (EPOC 2015) tool included in the Cochrane Handbook (Higgins & Green 2011). This tool addresses

participant allocation sequence; concealment of allocation; baseline outcome measurement; baseline participant characteristics; incomplete data reporting; blinding; data contamination; selective outcome reporting; and other biases. Furthermore, the risk of bias for pre- and post-test studies was assessed using seven criteria in relation to confounding variables; shape of the intervention; consistency in data collection at pre- and post-test; blinding; incomplete and/or selective outcome reporting; and other biases (EPOC 2015).

The quality of evidence for each review outcome (i.e. knowledge, confidence and clinical performance) was then assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool (Guyatt *et al.* 2008). This is a key step in systematic reviews of interventions, as failure to do so poses a threat to the accuracy of recommendations (Higgins & Green 2011, Saab *et al.* 2016b). The quality of evidence for each outcome was assessed in terms of limitations in the design and implementation; indirectness of evidence; heterogeneity or inconsistency of results; imprecision of results; and likelihood of publication bias (Guyatt *et al.* 2008). Accordingly, the overall quality (i.e. GRADE) of each outcome was rated as either high, moderate, low, or very low (Higgins & Green 2011).

Finally, the Scottish Intercollegiate Guidelines Network (SIGN 2014) level of evidence criteria were used to determine the level of evidence for each study in terms of internal validity (i.e. selection of subjects, assessment of outcomes, confounding and statistical analysis) and overall assessment. The level of evidence was graded between 1 and 4, with 1 being the highest score (SIGN 2014).

Data abstraction

Records identified from the search were exported to a software package for reference management (Endnote X7) and duplicates were deleted. All records were independently screened on title and abstract by the research team (in pairs) to determine whether they merited full-text review. Following the exclusion of ineligible articles based on title and abstract screening, the full-texts of potentially eligible papers were obtained and evaluated independently by two reviewers. Screening conflicts between two reviewers were resolved by consensus and if necessary involved a third reviewer.

Data from the included studies were extracted by two researchers using a standardised extraction table that was used in previous systematic reviews (Saab *et al.* 2016a; Saab *et al.* 2016b). The table was then cross-checked by two other reviewers for accuracy. The extracted data included: author(s) and year; aim(s); country and setting; population; study design; description of the intervention; outcomes measured; and findings. The level of evidence for each study using the SIGN (2014) criteria was also included in the data extraction table.

Synthesis

This review is reported using the items of the PRISMA checklist (Moher *et al.* 2009). Data extracted from the reviewed studies are tabulated (Table 1). The study selection process is then discussed in detail in terms of study identification, screening and inclusion. A synthesis of the key study characteristics (i.e. country and setting, population, design and outcomes measured) is then presented in-text and in a table format (Table 2). Findings from

individual studies are presented in the data extraction table. Narrative summaries of data were produced and grouped for each of the review outcomes (i.e. knowledge, confidence and clinical performance).

A meta-analysis was not attempted for several reasons. According to Higgins and Green (2011), a meta-analysis is not possible when the outcomes are varied and the studies are clinically diverse which is the case in this review. Moreover, studies in the present review differed in terms of methodology, educational programmes, modes of delivery, duration of programmes, instruments used to measure programme effectiveness, clinical settings where programmes were tested and length of follow-up.

The review outcomes were synthesized and mapped using a harvest plot (Figure 1) (Turley *et al.* 2013). Outcomes were plotted on the vertical axis. The direction of effects was plotted on the horizontal axis using three categories: 'favours control', 'no difference' and 'favours intervention'. Each of the reviewed studies was represented using a bar with the first three letters of the first author's last name. Shading of the bars corresponded to the statistical confidence in point estimate, height of the bar indicated the appropriateness of study design and the symbol over each bar indicated the risk of bias suing the EPOC (2015) criteria (Turley *et al.* 2013).

RESULTS

Study selection

The search strategy yielded 3,598 titles and abstracts. Duplicates were deleted (n=294) and irrelevant records were excluded based on title and abstract screening (n=3,304). Following a full-text review of 267 articles, 256 full-text papers were excluded as they focused on budget impact analyses of EWS, evaluations of the implementation of EWS and effectiveness and validation of EWS systems. Reference list checks did not yield any new articles. Therefore, a total of 10 studies in 11 papers met the review eligibility criteria and were included in this review. Findings from the searches at each stage of the review process are illustrated using the PRISMA flowchart (Moher *et al.* 2009) (Figure 2).

Study characteristics

The study characteristics are presented in Table 2. The same study was reported in Liaw *et al.* (2015a) and Liaw *et al.* (2015b). Six countries were represented across the reviewed studies with the greater numbers conducted in USA (n=3), (Lindsey & Jenkins, 2013; Ozekcin *et al.* 2015; Rose *et al.* 2015), Singapore (n=2) (Liaw *et al.* 2014, Liaw *et al.* 2015a, Liaw *et al.* 2015b), in hospitals (n=7) (Cahill *et al.* 2011, Ludikhuize *et al.* 2011, Shaddel *et al.* 2014, Kyriacos *et al.* 2015, Merriel *et al.* 2015, Ozekcin *et al.* 2015, Rose *et al.* 2015) and simulation settings (n=3) (Lindsey & Jenkins 2013, Liaw *et al.* 2014, Liaw *et al.* 2015a, Liaw *et al.* 2015b). Sample size varied between 19 (Shaddel *et al.* 2014) and 147 (Cahill *et al.* 2011) participants. Five studies used a pre- and post-test design (Cahill *et al.* 2011, Shaddel *et al.* 2014, Merriel *et al.* 2015, Ozekcin *et al.* 2015, Rose *et al.* 2015), four studies were RCTs (Lindsey & Jenkins 2013, Liaw *et al.* 2014, Liaw *et al.* 2015a, Liaw *et al.*

2015b, Kyriacos *et al.* 2015) and one study was quasi-experimental (Ludikhuize *et al.* 2011). As for the educational interventions used in the reviewed studies, they included: interactive programmes, namely Rescuing a Patient in Deteriorating Situations (RAPIDS) (Liaw *et al.* 2014) and e-RAPIDS (Liaw *et al.* 2015a, Liaw *et al.* 2015b); fully automated virtual patient simulation and facilitator-led mannequin-based simulation (Liaw *et al.* 2014); MEWS charts and the Cape Town MEWS training programme and manual (Kyriacos *et al.* 2015); educational sessions about a new observation chart (Cahill *et al.* 2011); code blue simulation and rapid response education (Lindsey & Jenkins 2013); a fictional deteriorating patient (Ludikhuize *et al.* 2011); an EWS training session (Merriel *et al.* 2015); an e-learning module and simulation (Ozekcin *et al.* 2015); and one-on-one and small group education about e-MEWS (Rose *et al.* 2015) and MEWS (Shaddel *et al.* 2014).

The effectiveness of the interventions was assessed using several researcher designed instruments namely: a knowledge questionnaire and performance tool (Liaw *et al.* 2015a, Liaw *et al.* 2015b); written tests (Kyriacos *et al.* 2015); multiple choice questionnaires (Lindsey & Jenkins 2013, Ozekcin *et al.* 2015); clinical observations (Ludikhuize *et al.* 2011); and chart reviews and audits (Merriel *et al.* 2015, Rose *et al.* 2015).

Risk of bias and level of evidence assessment

Nurses in the reviewed controlled trials were adequately allocated to control and experimental groups and their allocation was adequately concealed with the exception of one quasi-experimental study (Ludikhuize *et al.* 2011). As for baseline outcome measures, the risk of bias was low in two studies (Kyriacos *et al.* 2015, Liaw *et al.* 2015a, 2015b) and

unclear in the remaining studies (Ludikhuize *et al.* 2011, Lindsey & Jenkins 2013, Liaw *et al.* 2014). Participant characteristics were similar at baseline in all but one RCT (Linsey & Jenkins 2013). Incomplete outcomes were not addressed in the majority of the controlled trials. Blinding was reported in all but one RCT (Lindsey & Jenkins 2013) and the risk for data contamination was low in all controlled trials with the exception of one RCT (Lindsey & Jenkins 2013). All the reviewed controlled trials were free from selective outcome reporting, two studies were free from other risks of bias (Ludikhuize *et al.* 2011, Kyriakos *et al.* 2015) and estimates of precision were reported in only one RCT (Kyriakos *et al.* 2015). The full risk of bias assessment for RCTs and non-RCTs is available as a supplementary file (See supporting information Table S2 in the online version of the article).

As for the pre- and post-test studies, all but one intervention (Rose *et al.* 2015) were free from confounders. The shape of the intervention effect was pre-specified in all but one study (Shaddel *et al.* 2014). The data collected before and after the intervention were the same for all the reviewed pre- post-test studies. Blinding was addressed in two studies (Merriel *et al.* 2015, Rose *et al.* 2015) and omitted in the remaining three studies (Cahill *et al.* 2011, Shaddel *et al.* 2014, Ozekcin *et al.* 2015). The risk for incomplete outcome reporting was unclear in all five pre- and post-test studies. The risk for selective outcomes reporting was found to be low in all pre- and post-test studies. Yet, they had a high risk for other biases including data contamination. Moreover, only Merriel *et al.* (2015) reported on estimates of precision. The full risk of bias assessment for pre- and post-test studies is available as a supplementary file (See supporting information Table S3 in the online version of the article).

The overall quality of evidence rating for each outcome was found to be moderate for knowledge and low for performance and confidence. This was attributed to several methodological limitations and biases. For instance, not all the reviewed studies addressed blinding of the outcome assessor and only four studies had a robust design (i.e. RCT). In addition, the effectiveness of the reviewed educational interventions was often assessed using researcher-designed instruments with no details as to their validity or reliability. As for imprecision, most of the reviewed interventions had a small sample size that was selected purposely rather than randomly. Differences in baseline outcome measures and reporting on incomplete outcome data were unclear in the majority of the reviewed interventions. The quality of evidence assessment per review outcome is presented in Table 3.

In relation to the level of evidence for each study, all but one (Kyriacos *et al.* 2015) scored 2- on the SIGN tool which indicates a high risk of confounding, bias, as well as a significant risk that the relationship between the variables is not causal (SIGN 2014). The scores per study reviewed are presented in Table 1.

Synthesis of results

Knowledge

Knowledge was assessed in four studies (Lindsey & Jenkins 2013, Kyriacos *et al.* 2015, Liaw *et al.* 2015a, Liaw *et al.* 2015b, Ozekcin *et al.* 2015). Overall, the knowledge and competence of healthcare professionals improved immediately following various educational programmes. For instance, knowledge of the key elements of EWS significantly increased among nurses who attended an interactive web-based programme (e-RAPIDS) in comparison

to those who did not (21.29% vs. 18.89%; P<0.001) (Liaw *et al.* 2015a, Liaw *et al.* 2015b). Similar findings were reported by Lindsey and Jenkins (2013) whereby a novel rapid response education intervention succeeded in enhancing nursing students' understandings of rapid response systems compared with those who did not receive the education (Mean=90.91 SD 8.73 vs. 64.8 SD 19.69 respectively; P<0.001). In another study, Kyriacos *et al.* (2015) introduced a novel MEWS chart and associated training which was found to increase nurses' knowledge scores from a mean of 4/23 (19.5%) at pre-test to 14/23 (61.4%) (t3.8; 95%CI -30.0t, 8.9; P=0.001) two weeks following the intervention.

Confidence

Two of the reviewed studies measured the nurses' level of confidence (Shaddel *et al.* 2014, Ozekcin *et al.* 2015). For instance, Ozekcin *et al.* (2015) investigated the effectiveness of a four-week e-learning module on nurses' knowledge of signs and symptoms of deterioration and confidence in recognising clinical deterioration. It was found that, following the module, nurses' confidence increased significantly in recognising deterioration (Mean=4.06/5 SD 0.44 at pre-test vs. 4.45/5 SD 0.51 at post-test; P=0.001) and in responding to an unstable patient (Mean=4/5 SD 0.52 at pre-test vs. 4.48/5±0.51 at post-test; P<0.0001). Shaddel *et al.* (2014) also explored nurses' confidence following the introduction of the MEWS tool and associated training. It was found that confidence significantly improved from a mean of 3.73/5 at pre-test to 4.63/5 at post-test (Z=3.81; P=0.0001). The long-term effects of both interventions were not reported (Shaddel *et al.* 2014, Ozekcin *et al.* 2015).

Clinical performance

Clinical performance was assessed in all but one study (Lindsey & Jenkins 2013) and was judged in terms of accurate documentation of vital signs, accurate calculation of EWS and appropriate response to clinical deterioration.

Generally, nurses in the reviewed studies correctly calculated early warning scores (i.e. recorded a full set of vital signs and computed the corresponding EWS) following exposure to the educational programmes (Ludikhuize *et al.* 2011, Liaw *et al.* 2014, Shaddel *et al.* 2014, Liaw *et al.* 2015a, Liaw *et al.* 2015b, Merriel *et al.* 2015, Ozekcin *et al.* 2015). Merriel *et al.* (2015) evaluated the effectiveness of multidisciplinary training on intervention and recognition of the deteriorating patient. It was found that nurses were more likely to calculate early warning scores correctly post-test in comparison to pre-test (68.02% vs. 55.12%; Risk Ratio=1.24, 95% CI 1.07,1.44; P<0.01). In addition, observations were more likely to be performed at the correct frequency compared with pre-test (78.57% vs. 68.09%; Risk Ratio=1.20, 95% CI 1.09, 1.32). Another example is the study by Liaw *et al.* (2014) whereby nursing students' performance improved significantly immediately and 2.5 months following a fully automated virtual patient simulation (P<0.001) and a facilitator-led mannequin-based simulation (P<0.05).

Recording of vital signs improved in four studies (Cahill *et al.* 2011, Kyriacos *et al.* 2015, Ludikhuize *et al.* 2011, Liaw *et al.* 2014). For instance, following the introduction of a new observation chart and associated education, documentation of a full set of vital signs improved significantly (47.6% at pre-rest vs. 96.3% two weeks post-test vs. 96.4% three

months post-test; P<0.001) (Cahill *et al.* 2011). Similarly, documentation of respiratory rate (48.2% vs. 25%; P<0.05) and heart rate (74.3% vs. 37.5%; P<0.01) improved significantly among nurses who undertook an EWS web-based programme (RAPIDS) in comparison to those who did not (Liaw *et al.* 2015a, Liaw *et al.* 2015b).

MEWS training did not lead to an increase in appropriate response to clinical deterioration in several studies. For instance, although MEWS-trained nurses were able to identify and review a deteriorating patient more often than untrained nurses (77% vs. 58%; P=0.05), 67% of trained nurses and 43% of non-trained nurses notified the physician which was not statistically significant (Ludikhuize *et al.* 2011). In addition, only 11% of trained nurses calculated MEWS correctly and only 1 of 47 trained nurses used SBAR (situation, background, assessment and recommendation) (Ludikhuize *et al.* 2011). Similarly, Kyriacos *et al.* (2015) found that MEWS training was not associated with a significant change in response to deterioration among trained nurses (Odds Ratio=2.63; 95% CI 0.53, 12.97).

DISCUSSION

Evidence from this review suggests that EWS educational programmes succeeded in increasing nurses' knowledge, confidence and clinical performance with regards to calculation of EWS and documentation of vital signs, at least in the short-term (i.e. immediately following exposure to the programme). Several interventions had little or no effect on nurses' detection of clinical deterioration, appropriate escalation and use of communication tools such as SBAR. Examples include a study using MEWS chart, Cape Town MEWS training programme and manual (Kyriacos *et al.* 2015) and an observational study whereby MEWS trained nurses' responses to a fictional deteriorating patient was assessed (Ludikhuize *et al.* 2011).

This review confirms that there is lack of high quality evidence to evaluate the effect of EWS educational programmes on nurses' knowledge, confidence and clinical performance. This was thought to be due to several factors including the small sample size, lack of evidence of sample size calculation, lack of blinding of the outcome assessors and biases. Several publications relating to researcher-designed programmes lacked details as to the contents of the educational interventions. In addition, a variety of outcomes were measured using various tools and studies were heterogeneous in terms of methodology and clinical setting. This made it impossible to group the review outcomes into a meta-analysis (Higgins & Green 2011).

Time of delivery of the educational sessions varied enormously, from 15 minutes in one study (Shaddel *et al.* 2015) to 8 months in another (Liaw *et al.* 2014) with no study using well-established educational programmes such as AIM, ALERT, COMPASS and MFS (Liaw *et al.* 2011). For example, COMPASS is known to be effective in the categorisation of patients' severity of illness, early detection of patient deterioration, use of communication tools such as SBAR and the identification of triggers points that should prompt early medical review and use of an escalation plan (Health Service Executive 2011). However, the use of COMPASS alone does not guarantee that appropriate escalation of care is going to take place.

Although the key assessment parameters addressed in EWS were addressed in the reviewed educational programmes, other parameters that have shown to predict clinical deterioration and adverse outcomes were not accounted for. These include patient age (Churpek *et al.* 2015), urinary output (Martin *et al.* 2015), emotional state (Bian *et al.* 2015),

frailty (Romero-Brufau *et al.* 2014), diastolic blood pressure (Christofidis *et al.* 2013), pulse pressure index (Churpek *et al.* 2012), prior admission to ICU (Churpek *et al.* 2014) and pre-existing comorbidities (Huggan *et al.* 2015, Hegarty *et al.* 2016).

Measurements of knowledge, confidence and clinical performance varied across the reviewed studies. For instance, performance was judged on the basis of the frequency of vital signs monitoring; escalation and MEWS calculation; time to application of critical interventions; number of code blue and rapid response team calls; and the appropriateness of decisions regarding the management of deteriorating patients. As for knowledge and confidence, both outcomes were measured using clinical observations; researcher-designed tools; multiple choice questionnaires; and written tests with limited information as to their reliability and validity. It is worth noting that only Cahill *et al.* (2011) and Liaw *et al.* (2014) explored the longitudinal effect of the educational interventions (3 and 2.5 months post-education respectively). Therefore, the effectiveness of the reviewed interventions in increasing knowledge, confidence and clinical performance in the long-term remains unknown.

In relation to participants and data collection settings, the educational interventions were tested amongst nurses and nursing students with limited representation of other healthcare professionals. This undermines the important role of the healthcare team in the detection and management of clinical deterioration, given that there is evidence that interdisciplinary and multimodal educational programmes are effective in enhancing the use of EWS (Liaw *et al.* 2014, Hegarty *et al.* 2016). The majority of the reviewed studies were conducted either in simulation settings or in hospitals using fictitious patients which makes

their applicability to real-life scenarios questionable (Oberleitner *et al.* 2011). It is worth noting that all but one study (Kyriacos *et al.* 2015) were conducted in countries with very high human development index. Moreover, some studies comprised sample sizes as small as 19 (Shaddel *et al.* 2014), which hinders precision.

Findings from the reviewed studies demonstrated improved knowledge and confidence but only in the short-term. While there was some improvement in performance in relation to vital signs recording and EWS calculation, it remains unknown if this improvement is maintained over time and what effect is has on patient outcomes. Finally, the review highlights that education in isolation from other factors is not enough to enhance knowledge, confidence and clinical performance. This was evident in EWS trained nurses' failure to correctly respond to clinical deterioration on several occasions (Kyriacos *et al.* 2015; Ludikhuize *et al.* 2011).

Rigour was sought throughout the systematic review process by using the PRISMA checklist (Moher *et al.* 2009) in the reporting of this review and thoroughly describing study identification, screening, selection and data extraction. However, the search was limited to studies published in or translated to English between the years 2011 and 2015, thus increasing the risk of study selection bias. Furthermore, only findings that were in line with the review outcomes (i.e. knowledge, confidence and clinical performance) were extracted and discussed which increases the risk of reporting bias and could have contributed to the omission of potentially important findings (Cochrane Bias Methods Group 2013).

Future research is needed to address the limitations highlighted in this systematic review. Researchers are encouraged to explore, in-depth, the reason why several interventions did not have an effect on nurses' detection of clinical deterioration and escalation of care. In addition, researchers ought to conduct studies with larger sample sizes and use measures to minimise bias including blinding the outcome assessor, random sampling and controlling for possible confounders (e.g. level of nurses' autonomy and power relationships between nurses and physicians). Researchers are also encouraged to provide more details as to the content of the educational programmes and to test well-established programmes such as AIM, ALERT, COMPASS and MFS. It is also worth accounting for other assessment parameters which can influence clinical judgment, including patient age, urinary output, emotional state, frailty, diastolic blood pressure, pulse pressure index, prior admission to ICU and pre-existing comorbidities.

Longitudinal studies are needed to explore the long-term effect of the educational interventions on nurses' knowledge, confidence and clinical performance while using valid and reliable instruments. Finally, researchers are encouraged to systematically review the evidence on the effect of EWS educational programmes on patient outcomes including mortality, ICU transfers and length of hospital stay.

CONCLUSION

There is lack of high quality evidence to evaluate the effect of EWS educational programmes on nurses' knowledge, confidence and clinical performance. Given that EWS themselves represent a complex intervention this can only be achieved by using techniques that go beyond enhancing knowledge, confidence and move towards consistent clinical performance in the real world. This mandates the need for valid and reliable outcome-based

training programmes, which deploy several approaches including interactive e-learning, workshops and practice in the clinical setting. Finally, it is likely that effectiveness will be enhanced if educational interventions are interdisciplinary, delivered frequently and measured longitudinally.

Author Contributions:

All authors have agreed on the final version and meet at least one of the following criteria (recommended by the ICMJE*):

1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data;

2) drafting the article or revising it critically for important intellectual content.

* http://www.icmje.org/recommendations/

REFERENCES

Alam N, Vegting IL, Houben E, et al. (2015) Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation, 90, 111-115.

Bian Y, Xu F, Lv R-J, et al. (2015) An early warning scoring system for the prevention of acute heart failure. Int J Cardiol 183, 111-116.

Cahill H, Jones A, Herkes R, et al. (2011) Introduction of a new observation chart and education programme is associated with higher rates of vital sign ascertainment in hospital wards. BMJ Qual Saf, 20, 791-796.

Christensen D, Jensen NM, Maaløe R, Rudolph SS, Belhage B, Perrild H. (2011) Nurse-administered early warning score system can be used for emergency department triage. Dan Med Bull, 58, A4221-A4221.

Christofidis MJ, Hill A, Horswill MS, Watson MO (2013) A human factors approach to observation chart design can trump health professionals' prior chart experience. Resuscitation, 84, 657-665.

Churpek MM, Yuen TC, Huber MT, Park SY, Hall JB, Edelson DP. (2012) Predicting cardiac arrest on the wards: a nested case-control study. CHEST, 141, 1170-1176.

Churpek MM, Yuen TC, Winslow C, et al. (2014) Multicenter development and validation of a risk stratification tool for ward patients. Am J Respir Crit Care Med 190, 649-655.

Churpek MM, Yuen TC, Winslow C, Hall J, Edelson DP. (2015) Differences in Vital Signs Between Elderly and Nonelderly Patients Prior to Ward Cardiac Arrest. Crit Care Med, 43, 816-822.

Cochrane Bias Methods Group (2013). Assessing risk of bias in included studies. (Accessed 11 April 2016, at https://bmg.cochrane.org/assessing-risk-bias-included-studies).

De meester K, Verspuy M, Monsieurs KG, van Bogaert P. (2013) SBAR improves nurse-physician communication and reduces unexpected death: a pre and post intervention study. Resuscitation, 84, 1192-1196.

Department of Health (2015) Publications. Retrieved from https://www.gov.uk/government/publications on 29 November 2015.

Effective Practice and Organisation of Care (2015) Suggested Risk of Bias Criteria for EPOC Reviews. Retrieved from

https://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/Suggested%20risk%20of%20bias%20criteria%20for%20EPOC%20reviews.pdf on 31 December 2016

Fagan K, Sabel A, Mehler PS, MacKenzie TD. (2012) Vital sign abnormalities, rapid response and adverse outcomes in hospitalized patients. Am J Med Qual, 27, 480-486.

Google (2015) Google. Retrieved from https://www.google.com/ on 29 November 2015.

Google Scholar (2015) Google Scholar. Retrieved from https://scholar.google.com/ on 29 November 2015.

Guyatt GH, Oxman AD, Vist GE, et al. (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ, 336, 924-926.

Harrison G, Jacques T, Kilborn G, McLaws ML. (2005) The prevalence of recordings of the signs of critical conditions and emergency responses in hospital wards - The SOCCER study. Resuscitation, 65, 149-157.

Health Information and Quality Authority (2015) Resource Centre. Retrieved from https://www.hiqa.ie/resource-centre on 28 November 2015.

Health Research Board (2015) Publications. Retrieved from http://www.hrb.ie/publications/ on 28 November 2015.

Health Service Executive (2011). Training manual for The National Early Warning Score and associated education programme (Accessed 20 April 2016, at http://www.hse.ie/eng/about/Who/clinical/natclinprog/acutemedicineprogramme/earlywarningscore/compass.pdf).

Health Service Executive (2015) HSE Reports and Publications. Retrieved from http://www.hse.ie/eng/services/publications/ on 30 November 2015.

Hegarty J, Drummond FJ, Murphy A, et al. (2016) A systematic review of the clinical & economic literature and a budget impact analysis of any new guideline recommendations

to inform the planned update of National Clinical Guideline No. 1 - National Early Warning Score (NEWS) for the Irish health system (Accessed 25 April 2016, at http://health.gov.ie/patient-safety/ncec/national-clinical-guidelines-2/).

Higgins JPT, Green S (editors) (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1. The Cochrane Collaboration (Accessed 25 May 2016, at www.cochrane-handbook.org).

Huggan PJ, Akram F, Er BHD, et al. (2015) Measures of acute physiology, comorbidity and functional status to differentiate illness severity and length of stay among acute general medical admissions: a prospective cohort study. Intern Med, 45, 732-740.

Jamieson E, Ferrell C, Rutledge DN. (2008) Medical emergency team implementation: experiences of a mentor hospital. Medsurg Nurs, 17, 312-6, 323.

Kyriacos U, Jelsma J, James M, Jordan S. (2015) Early warning scoring systems versus standard observations charts for wards in South Africa: a cluster randomized controlled trial. Trials, 6. (Accessed 10 April 2016, at doi:10.1186/s13063-015-0624-2).

Lenus (2015) Browsing Lenus, The Irish Health Repository. Retrieved from http://www.lenus.ie/hse/browse on 28 November 2015.

Levett-Jones T, Hoffman K, Dempsey J, et al. (2010) The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. Nurse Educ Today, 30, 525-520.

Liaw SY, Chan SW-C, Chen F-G, Hooi SC, Siau C. (2014) Comparison of virtual patient simulation with mannequin-based simulation for improving clinical performances in assessing and managing clinical deterioration: randomized controlled trial. J Med Internet Res, 16, e214-e214.

Liaw SY, Scherpbier A, Klainin-Yobas P, Rethans JJ. (2011) A review of educational strategies to improve nurses' roles in recognizing and responding to deteriorating patients. Int Nurs Rev, 58, 296-303.

Liaw SY, Wong LF, Ang SBL, Ho JTY, Siau C, Ang ENK. (2015a) Strengthening the afferent limb of rapid response systems: an educational intervention using web-based learning for early recognition and responding to deteriorating patients. BMJ, 0, 1-9.

Liaw SY, Wong LF, Chan SW-C, et al. (2015b) Designing and evaluating an interactive multimedia web-based simulation for developing nurses' competencies in acute nursing care: randomized controlled trial. J Med Internet Res, 17, e5.

Lindsey PL, Jenkins S. (2013) Nursing students' clinical judgment regarding rapid response: the influence of a clinical simulation education intervention. Nurs Forum, 48, 61-70.

Ludikhuize J, De Jonge E, Goossens A. (2011) Measuring adherence among nurses one year after training in applying the Modified Early Warning Score and Situation-Background-Assessment-Recommendation instruments. Resuscitation, 82, 1428-1433.

Ludikhuize J, Smorenburg SM, De Rooij SE, De Jonge E. (2012) Identification of deteriorating patients on general wards; measurement of vital parameters and potential effectiveness of the Modified Early Warning Score. Journal Of Critical Care, 27, 424.e7-13.

Martin G, Dupré A, Mulliez A, Prunel F, Slim K, Pezet D. (2015) Validation of a score for the early diagnosis of anastomotic leakage following elective colorectal surgery. J Visc Surg, 152, 5-10.

Merriel A, van der Nelson H, Merriel S, et al. (2015) Identifying deteriorating patients through multidisciplinary team training. Am J Med Qual (Accessed 5 April 2016, at doi:10.1177/1062860615598573).

Mitchell I, McKay H, van Leuvan C, et al. (2010) A prospective controlled trial of the effect of a multi-faceted intervention on early recognition and intervention in deteriorating hospital patients. Resuscitation, 81, 658-666.

Moher D, Liberati A, Tetzlaff J, Altman DG. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med, 6, 1-6.

National Clinical Effectiveness Committee (2013). National Early Warning Score, national clinical guideline no.1 (Accessed 4 April 2016, at http://health.gov.ie/wp-content/uploads/2015/01/NEWSFull-ReportAugust2014.pdf).

National Health Service England (2015) Publications. Retrieved from https://www.england.nhs.uk/category/publications/ on 29 November 2015.

National Institute for Health and Care Excellence (2015) Find guidance. Retrieved from https://www.nice.org.uk/guidance on 29 November 2015.

New York Academy of Medicine (2015) Library. Retrieved from http://nyam.org/library/ on 28 November 2015.

Oberleitner MG, Broussard AB, Bourque J. (2011) An unintended consequence of simulation: a case report. Clin Simul Nurs, 7, e35–e40.

Open Grey (2015) System for Information on Grey Literature in Europe. Retrieved from http://www.opengrey.eu/ on 28 November 2015.

Ozekcin LR, Tuite P, Willner K, Hravnak M. (2015) Simulation education: early identification of patient physiologic deterioration by acute care nurses. Clin Nurse Spec, 29, 166-173.

Public Health Agency of Canada (2015) Reports & Publications. Retrieved from http://www.phac-aspc.gc.ca/publications-eng.php on 30 November 2015.

Rabol LI andersen ML, Ostergaard D, Bjorn B, Lilja B, Mogensen T. (2011) Descriptions of verbal communication errors between staff. An analysis of 84 root cause analysis-reports from Danish hospitals. BMJ Qual Saf, 20, 268-274.

Robb G, Seddon MA (2010) multi-faceted approach to the physiologically unstable patient. Qual Saf Health Care, 19, e47.

Romero-Brufau S, Huddleston JM, Naessens JM, et al. (2014) Widely used track and trigger scores: Are they ready for automation in practice? Resuscitation, 85, 549-552.

Rose MA, Hanna LA, Nur SA, Johnson CM. (2015) Utilization of electronic modified early warning score to engage rapid response team early in clinical deterioration. J Nurses Prof Dev 31, E1-E7.

Saab MM, Landers M, Hegarty J (2016a) Promoting testicular cancer awareness and screening: a systematic review. Cancer Nursing, 39(6), 473-787.

Saab MM, Landers M, Hegarty J (2016b) Testicular cancer awareness and screening practices: a systematic review. Oncology Nursing Forum, 4391), E8-E23.

Scottish Intercollegiate Guidelines Network (2014). Critical appraisal: notes and checklists (Accessed 10 April 2016, at http://www.sign.ac.uk/methodology/checklists.html).

Shaddel F, Khosla V, Banerjee S. (2014) Effects of introducing MEWS on nursing staff in mental health inpatient settings. Prog Neurol Psychiatry, 18, 24-27.

Shearer B, Marshall S, Buist MD, et al. (2012) What stops hospital clinical staff from following protocols? An analysis of the incidence and factors behind the failure of bedside clinical staff to activate the rapid response system in a multi-campus Australian metropolitan healthcare. BMJ Qual Saf 21, 569-575.

Subbe CP, Kruger M, Rutherford P, Gemmel L. (2011) Validation of a modified Early Warning Score in medical admissions. QJM, 94, 521-526.

Taenzer AH, Pyke JB, McGrath SP. (2011) A review of current and emerging approaches to address failure-to-rescue. Anesthesiology, 115, 421-431.

Turley R, Saith R, Bhan N, Rehfuess E, Carter B. (2013) Slum upgrading strategies involving physical environment and infrastructure interventions and their effects on health and socio-economic outcomes. Cochrane Database Syst Rev, 1(1), CD010067.

Urban RW, Mumba M, Martin SD, Glowicz J, Cipher DJ. (2015) Modified Early Warning System as a predictor for hospital admissions and previous visits in emergency departments. Adv Emerg Nurs J, 37, 281-289.

World Health Organization (2015) Publication. Retrieved from http://www.who.int/publications/en/ on 29 November 2015.

Table 1
Study characteristics, findings, and level of evidence assessment (Saab *et al.* 2016a, Saab *et al.* 2016b).

Author(s) (Year)	Aim(s)	Country & Setting	Population	Design	Intervention	Outcomes Measured	Findings	SIGN[a]
Cahill *et al.* (2011)	To evaluate the impact of a new observation chart and education on vital sign recording	Australia Hospital (3 medical/ surgical wards)	n=104 (T1) n=147 (T2) n=119 (T3) Patients	Prospective pre- and post-test	New observation chart and educational programme	Performance (vital sign recording)	Documentation of full vital signs increased significantly (47.6% at T1 vs. 96.3% at T2 vs. 96.4% at T3; P<0.001)	2-
Kyriacos *et al*. (2015)	To test the impact of a new MEWS chart and training on nurses' responses to clinical deterioration	South Africa Hospital (6 surgical wards)	n=50 Nurses	Pragmatic, parallel-group, cluster RCT	EG: MEWS charts and Cape Town MEWS training programme and manual CG: standard care	Knowledge (signs of deterioration) Performance (vital sign recording; response to deterioration)	Increased significantly among EG between T1 (Mean=4/23; 19.5%) and T2 (Mean=14/23; 61.4%) (t3.8; 95%CI -30.0t, 8.9; P=0.001); Increase was not significant among CG Vital signs recorded among EG > CG Unrecorded responses to MEWS triggers: 94.5% for EG and 97.8% for CG (OR=2.63; 95%CI 0.53, 2.97) No significant change in response to deterioration among EG	2+
Liaw *et al.* (2014)	To assess the impact of a new automated virtual patient simulation versus mannequin-based	Singapore Simulation	n=57 (EG=31; CG=26) Nursing	RCT; pre- and post-test	EG and CG: RAPIDS[b] simulation course eight months earlier EG: automated virtual	Performance (assessment; management; reporting of deterioration)	Increased significantly among EG (P<0.001) and CG (P<0.05) from T1 to T2 and T3 No significant difference between T2 and	2-

	simulation		students		patient simulation CG: mannequin-based simulation		T3 for EG and CG	
Liaw *et al.* (2015a; 2015b)[c]	To evaluate the impact of web-based simulation on the recognition of and response to patient deterioration	Singapore Simulation	n=67 (EG=32; CG=32) Nurses	RCT; pre-and post-test	EG: interactive web-based programme (e-RAPIDS[d]) CG: no intervention	Knowledge (assessment; management; communication of deterioration Performance (vital-sign recording; assessment; management; reporting of deterioration)	Significantly higher among EG at T2 compared to T1 (21.29% vs. 18.89%; P<0.001) Significantly higher among EG compared to CG (F=25.26; P<0.001) RR (48.2% vs. 25%; P<0.05) and HR measurements (74.3% vs. 37.5%; P<0.01) significantly higher among EG compared to CG; Positive correlation between assessment and clinical judgment (r=0.6, P<0.001) and reporting and assessment (r=0.56, P<0.001) among EG at T2	2-
Lindsey & Jenkins (2013)	To explore the impact of an intervention on students' clinical judgment in relation to managing patient deterioration	USA Simulation	n=79 (EG=40; CG=39) Nursing students	RCT; pre- and post-test	EG and CG: code blue simulation EG: rapid response education	Knowledge (RRS)	EG had significantly higher scores than CG (Mean=90.91±8.73 vs. 64.8±19.69; P<0.001) at T2	2-
Ludikhuize *et al.*	To evaluate whether MEWS trained	Netherlands	n=95	Quasi-	Observation of assessments and	Performance (vital-sign	More trained (77%) than untrained nurses (58%) reviewed the patient immediately	2-

(2011)	nurses were more likely to recognize patient deterioration than untrained nurses	Hospital (3 medical and 3 surgical wards)	MEWS trained nurses (n=47) and untrained nurses (n=48)	experimental	responses of nurses to a fictional deteriorating patient	recording; response; communication of deterioration)	(P=0.05) Trained nurses recorded RR more often than untrained nurses (53% vs. 25%; $\chi 2=5.038$; P=0.025) No differences between the two groups in the measurement of other parameters 11% of trained nurses calculated MEWS correctly and only one nurse used SBAR	
Merriel *et al*. (2015)	To assess whether an EWS training intervention can improve the recognition of patient deterioration	UK Hospital (3 surgical wards)	n=102 Nursing staff (n=83) and junior doctors (n=19)	Observational; pre- and post-test	EWS training session using real-life scenarios, simple tools, and debriefing 250 patient charts randomly assessed	Performance (vital-sign recording; EWS calculation)	Participants were more likely to calculate EWS correctly at T2 compared to T1 (68.02% vs. 55.12%; Risk Ratio=1.24; 95%CI 1.07, 1.44; P<0.01) Observations at T2 were more likely to be performed at the correct frequency compared to T1 (78.57% vs. 68.09%; Risk Ratio=1.20; 95%CI 1.09, 1.32)	2+
Ozekcin *et al*. (2015)	To improve nurses' ability to assess deteriorating patients, recognize signs of deterioration, and escalate care	USA Hospital (cardiac surgery unit)	n=35 Nurses	Observational; pre- and post-test	E-learning module and simulation over 4 weeks	Knowledge (signs; communication) Confidence (recognition; escalation of care)	Increased significantly at T2 (84.6%) compared to T1 (56.9%) (P<0.0001) Increased significantly in recognising deterioration (Mean=4.06±0.44 at T1 vs. 4.45±0.51 at T2; P=0.001) and responding	2-

							to an unstable patient (Mean=4±0.52 at T1 vs. 4.48±0.51 at T2; P<0.0001)	
						Performance (time to intervention)	Time to application of first correct critical intervention was faster, decreasing from 37% to 25% between scenarios	
Rose *et al.* (2015)	To re-educate clinical caregivers in the use of eMEWS and engagement of the RRT	USA Hospital (3 community units)	n=108 Nurses (87 RNs; 9 nurse technologists; 8 nurse assistants; 3 practical nurses; 1 respiratory therapist)	Observational; pre- and post-test	One-on-one or small group education on eMEWS, recording and engaging RRT Self-evaluation of knowledge Retrospective audit of RRT and code blue during 90-day pre- and post- education	Performance (eMEWS documentation; RRT calls; code blue calls)	RRT calls decreased at T2 (17/90 days) compared to T1 (23/90 days; 0 deaths) 23 RRT calls (11 events) had undocumented eMEWS scores at T1 vs. no undocumented eMEWS at T2 Code blue calls decreased at T2 (1/90 days) compared to T1 (6/90 days; 1 death) eMEWS score range increased at T2 (Mean=3.2±1.79; range 1-6) compared to T1 (Mean=2.3±1.79; range 0-6)	2-
Shaddel *et al.* (2014)	To explore nurses' confidence and ability to make correct clinical decisions regarding patient deterioration	UK Hospital (1 learning disability unit and 2 forensic units)	n=19 Nurses	Survey; pre- and post-test	Education on MEWS via case studies and training Confidence measured	Confidence (soundness of judgment) Performance (management of deterioration)	Improved between T1 (Mean=3.73/5) and T2 (Mean=4.63/5; Z=3.81; P=0.0001) Correct decision regarding patient management increased significantly from 42.1% at T1 to 92.1% at T2 (P<0.00001)	2-

[a] Empirical literature characterised according to the SIGN level of evidence criteria (SIGN 2014)

[b] Mannequin-based simulation programme with two areas: assessing ABCDE and using SBAR.

[c] Same study reported in two papers.

[d] Interactive web-based programme with three areas: detecting changes in vital signs, assessing ABCDE, and using ISBAR to report clinical deterioration.

Abbreviations: ABCDE: airway, breathing, circulation, disability, exposure; BP: blood pressure; CG: control group; CI: confidence interval; EG: experimental group; EWS: Early Warning Score; HR: heart rate; ISBAR: identify, situation, background, assessment, recommendation; MEWS: Modified Early Warning Score; OR: odds ratio; RAPIDS: rescuing a patient in deteriorating situations; RCT: randomized controlled trial; RN: registered nurse; RR: respiratory rate; RRS: rapid response system; RRT: rapid response team; SBAR: situation, background, assessment, recommendation; SIGN: Scottish Intercollegiate Guidelines Network; T1: pre-test; T2: post-test; T3: second post-test (follow-up).

Table 2

Key study characteristics (n=10 studies in 11 papers).

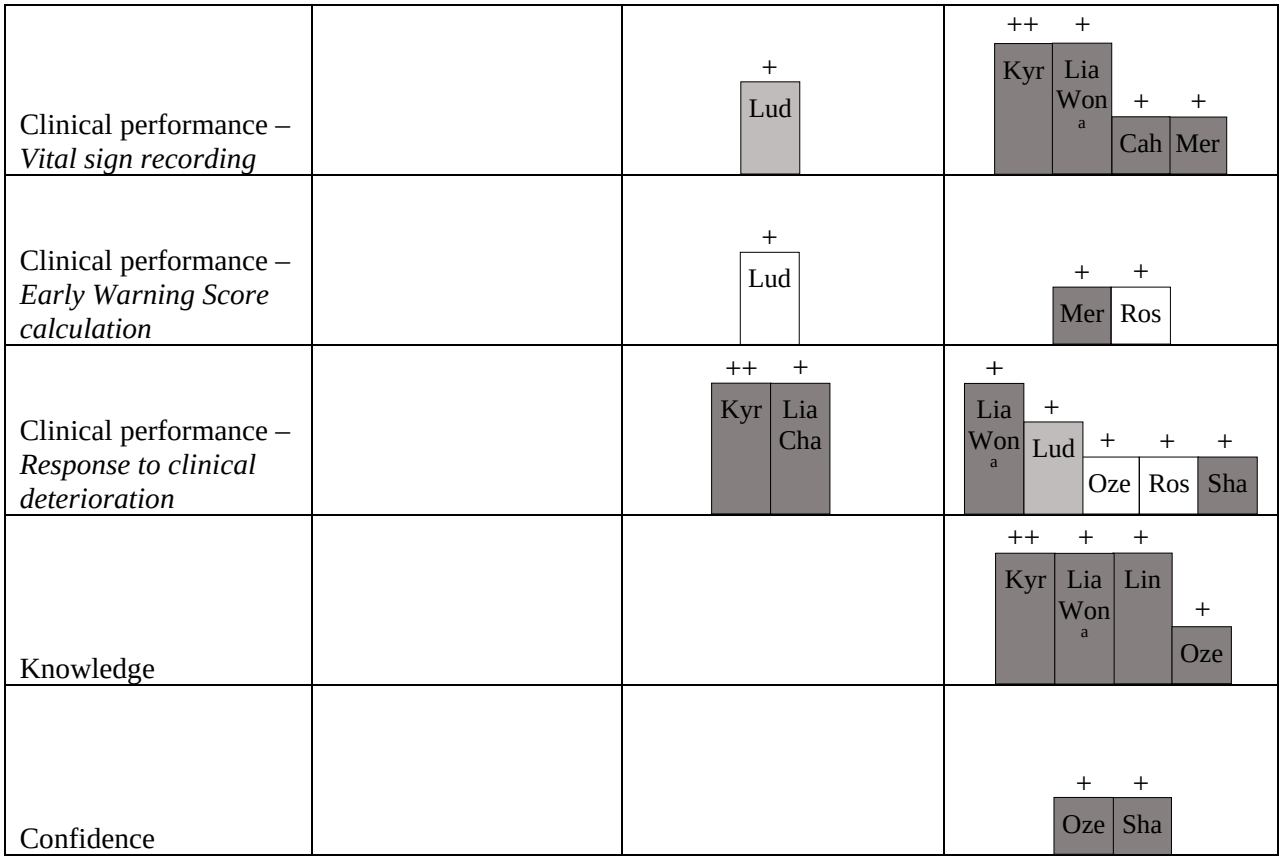
Country	USA (n=3)
	Singapore (n=2)
	UK (n=2)
	Australia (n=1)
	Netherlands (n=1)
	South Africa (n=1)
Setting	Hospital (n=7)
	Simulation (n=3)
Sample Size (min-max)	19-147
Study Design	Pre- and post-test (n=5)
	Randomised controlled trial (n=4)
	Quasi-experimental (n=1)
Outcomes Measured	Clinical performance (n=9)
	Knowledge (n=4)
	Confidence (n=2)

Table 3

Quality of evidence assessment per review outcome (Guyatt *et al.* 2008).

Outcomes	Number of participants (Number of studies) Follow-up	Limitations in the design and implementation	Indirectness of evidence	Unexplained heterogeneity or inconsistency	Imprecision of results	High probability of publication bias	Overall quality (GRADE)
Knowledge	231 (4 studies) NR	No	No	No	Yes	No	+++O Moderate
Confidence	54 (2 studies) NR	Yes	Yes	No	No	No	++OO Low
Performance	680 (9 studies) 2.5–3 months	Yes	Yes	No	No	No	++OO Low

Abbreviations: NR: not reported



[a] Same study reported in two papers.

KEY

Each bar corresponds to one study using the first three letters of the first author's family name. When two authors have the same family name, the first three letters of the second author's family name are also used. Study characteristics are represented as follows:

Shading of bar indicating the statistical confidence in point estate

Evidence of no effect or statistically significant effect at 1% level
Statistically significant effect at 5% level
Confidence intervals and p-values not reported/estimable

Height of bar indicating the appropriateness of the study design

High bar: design examining causal effect of intervention (RCT)
Medium bar: design inferring plausible causality (controlled before-after [CBA]/controlled post-intervention [CPI] with matching)
Low bar: design cannot examine causality (CBA/CPI)

Symbol indicating risk of bias per study using the EPOC (2015) criteria

++ Low risk of bias
+ Mixed/unclear risk of bias
– High risk of bias

Figure 1. Harvest plot synthesizing results from the reviewed studies (Turley *et al.* 2013)

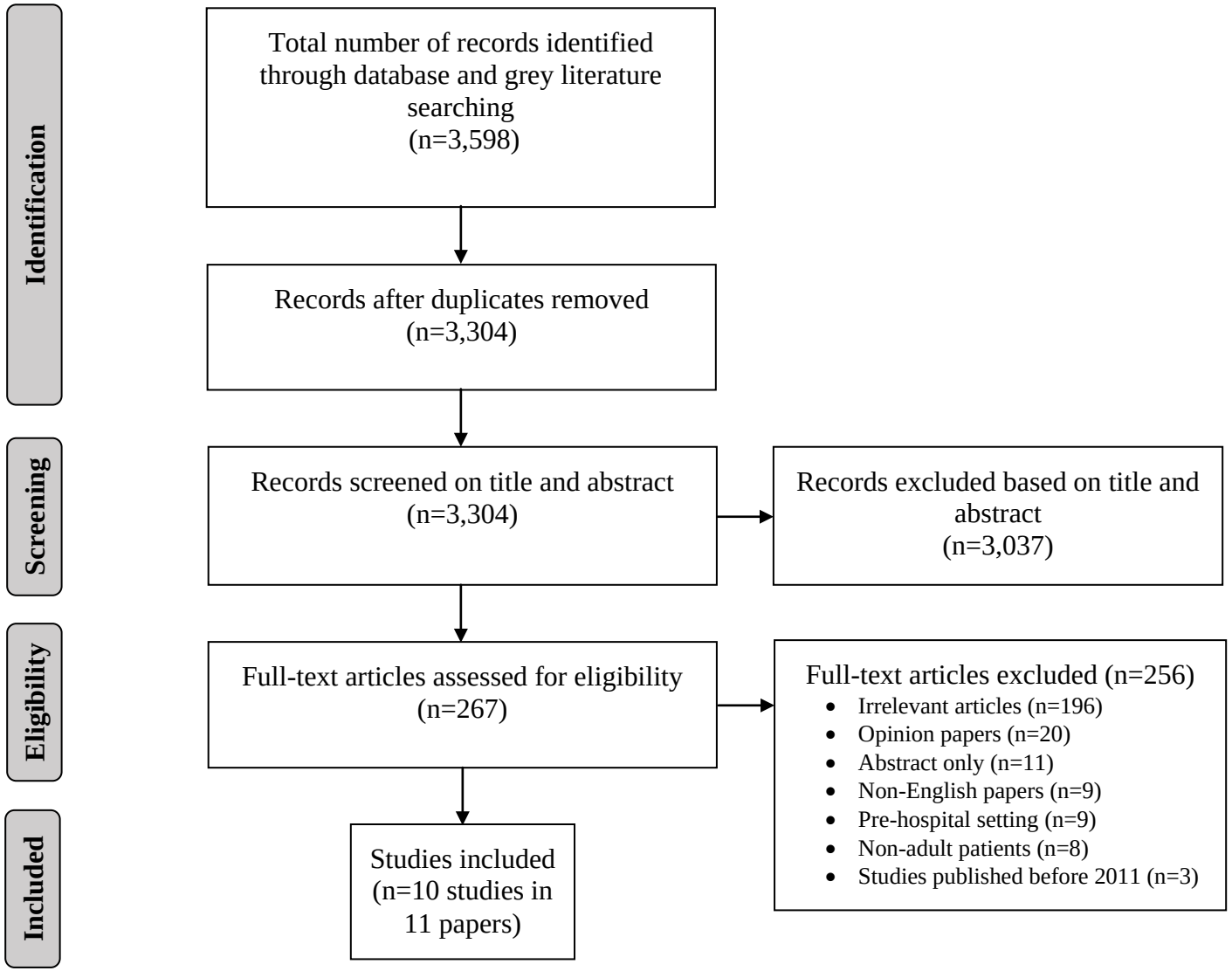


Figure 2. PRISMA flowchart (Moher *et al.* 2009)