CRITICAL APPRAISAL: ADEQUACY OF REPORTING STUDIES ON EARLY WARNING SCORE SYSTEMS

U Kyriacos
Division of Nursing & Midwifery, University of Cape Town

Abstract

Background: Routine monitoring of patients' clinical and physiological status by nurses includes the use of vital signs (observations) charts for recording findings. Charts that incorporate early warning score (EWS) systems are designed to 'track' signs of deterioration and 'trigger' a rapid response. Published studies of EWS systems are of limited benefit if reporting of these studies is inadequate. Reporting guidelines are recommended to improve the quality of reporting.

Objective: To assess the adequacy of reporting studies on EWS systems.

Methods: All study designs published between 1 May 2007 and 23 May 2015 describing the use of EWS systems for detecting deterioration in adult patients in general medical and surgical wards were included. Data extraction was undertaken by one researcher.

Results: Of the 657 references identified from search terms, 596 articles were excluded leaving 61 articles for assessment. Most studies were published in non-nursing journals (47/61, 77.0%). Six of the 61 (9.8%) studies were reviews. The remaining 55 clinical studies on the use of EWS and Modified EWS (MEWS) systems were mostly observational (46/55, 83.6%) rather than experimental (9/55, 16.4%).

Reporting guidelines were used in 9.8% (6/61) of reviewed studies. Only the reviews but no clinical studies reported a search strategy. Electronic searches included mostly CINAHL (5/6, 83.3%), MEDLINE and The Cochrane Library (4/6, 66.7%). No meta-analyses were performed. Inclusion and exclusion criteria and reasons for exclusion of references were well reported in the reviews.

The most frequently reported range of physiological parameters (12/61, 19.7%) were respiratory rate, oxygen saturation, supplemental oxygen, heart rate, systolic blood pressure, temperature and level of consciousness.

Conclusion: Reporting of published studies on EWS systems reviewed for this critical appraisal, with the exception of reviews, was inadequate as most did not use reporting guidelines, limiting the use of study findings for developing clinical guidelines and in further research.

Keywords: adult patients, deterioration, early warning score, general wards, modified early warning score, observations, vital signs.

INTRODUCTION

Nurses in general wards are responsible for the monitoring of patients' vital signs and escalation of care to high dependence care levels. However, acutely ill patients are increasingly being nursed in general wards where it is reported that vital signs' monitoring is infrequent and inadequate (Johnstone et al., 2007, Zimlichman et al., 2009). Nurses' responses to patients'clinical deterioration are reportedly also inappropriate (NHS NPSA, 2007) and delayed (Calzavacca et al., 2010, Calzavacca et al., 2008) despite changes in respiratory rates occurring six (Subbe et al., 2003) to eight hours before cardiopulmonary arrest. The delays are happening even though early interventions have been found to improve patient outcomes (Cioffi, 2000a).

According to Reason (2000), 70-80% of adverse events (AEs) in complex health care systems may be due to human error. Error may be reduced or curtailed by implementing structures for recognition of early warning signs with regard to clinical and physiological deterioration and appropriate algorithms for response (Wilson et al., 1999). Observation charts that incorporate early warning or modified early warning scoring (EWS/MEWS) systems (Kyriacos et al., 2014) and callout algorithms are bedside score and track-and-trigger systems: a total score is calculated to facilitate early recognition of a patient's deterioration.

In the UK a standardized National Early Warning Score (NEWS) system has been advocated (Royal College of Physicians, 2012) for monitoring six parameters (respiratory rate, oxygen saturations, temperature, systolic blood pressure, heart rate and level of consciousness) to improve patient outcomes. Having a system of 'tracking' early clinical and physiological deterioration in a patient and 'triggering' a predetermined reporting algorithm by specially trained nurses ought to benefit patients as this should improve patient safety by reducing the incidence of in-hospital deaths.

EWS/MEWS systems are used in conjunction with nurses' clinical judgement. Clinical signs such as skin tone, sweating, nausea or nurses' intuitive assessment of the patient being 'just not right' and 'looking unwell' (Cioffi, 2000b) should be monitored regularly to limit avoidable, serious adverse events (SAEs) such as cardiac arrest, urgent and unanticipated admission to an intensive care unit (ICU) or even death. In addition to obvious ethical considerations, authorities in the developed world are concerned at the increasing number of claims for malpractice associated with SAEs (Mello et al., 2003).

There are published observational studies (NHS NPSA, 2007, Wilson et al., 1999) and before and after evaluation studies (Reason, 2000) but only one randomized controlled trial (RCT) (Kyriacos et al., 2015) on the implementation and evaluation of MEWS training programs and recording systems. Published studies and reviews of EWS systems are of limited benefit if reporting of these studies is inadequate.

By 1996 there were concerns about the suboptimal reporting of published studies in general and of meta-analyses of RCTs in particular (Moher et al., 2009), limiting the use of study findings for developing clinical guidelines and in further research. In 1996 an international group developed guidelines, the QUOROM Statement (QUality Of Reporting Of Meta-analyses), focussing on the reporting of meta-analyses of RCTs (Moher et al., 1999). In 2005 these guidelines were revised and renamed PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) to improve reporting of meta-analyses and systematic reviews (Liberati et al., 2009). Reporting guidelines

enable researchers to report key aspects of research studies accurately and fully. In 2008, the EQUATOR (Enhancing the QUAlity and Transparency of Health Research) Network was launched (EQUATOR, 2013) to support wider practical implementation of reporting guidelines to increase the ease of use and value of health research for many different study designs.

Peer reviewers are encouraged to use reporting guidelines to enhance the utilisation of studies in clinical practice and in further research (Hirst and Altman, 2012). Examples of reporting guidelines for various types of studies (EQUATOR, 2013) are presented in Table 1.

Table 1: Reporting guidelines for main study types

Study type	Reporting guideline	Modifications
Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Case reports	CARE	-
Qualitative research	SRQR	COREQ
Diagnostic/prognostic studies	STARD	TRIPOD
Quality improvement studies	SQUIRE	-
Economic evaluations	CHEERS	-
Animal pre-clinical studies	ARRIVE	-
Study protocols	SPIRIT	PRISMA-P

(EQUATOR Network)

This article is a critical appraisal of the adequacy of reporting of all study designs and reviews describing the use of EWS and modified EWS (MEWS) systems for detecting deterioration in adult patients in general medical and surgical wards using full text publications in English from 1 May 2007 to 23 May 2015. The search strategy was guided by keywords to identify relevant sources. The structure of the paper was guided by the published literature on critical appraisal of articles (Dechartres et al., 2011, Monaghan, 2015).

METHODS

Searching for relevant articles

Electronic and additional searches are described as well as data extraction and statistical analysis.

Electronic search

An extensive electronic search was conducted by a single reviewer (Monaghan, 2015) using Cinahl (Cumulative Index of Nursing and Allied Health Literature) from the EBSCOHost database, PubMed and The Cochrane Database to locate articles published between 1 May 2007 and 23 May 2015. Keywords were limited to: vital signs or observations; early warning score; modified early warning score; deterioration; adult patients and general wards. The start date coincided with the end date of a systematic review of the literature on the nurses' role in detecting deterioration in ward patients conducted between 1992 and April 2007 (Odell et al., 2009).

For the PubMed search there were no MeSH terms corresponding to articles assessing the adequacy of published studies on EWS systems. Therefore all articles that met the inclusion criteria (Table 2) were searched using the seven keywords. Limits were set for the PubMed search and for Cinahl

(EBSCOHost) and articles were not included if they were not in English, not peer reviewed, if no full text was available, if they were published before 2007, if they were in the format of a report, a response, editorial or thesis, if they were duplicates and if exclusion criteria were met (Table 2). For the Cochrane search, the Database of Systematic Reviews was searched (Issue 1 of 12, January 2016) using the seven keywords.

Search strategy inclusion and exclusion criteria

Inclusion and exclusion criteria for the search strategy are presented in Table 2.

Table 2: Inclusion and exclusion criteria

Search inclusion criteria	Search exclusion criteria
• Research describing nursing observations (vital signs) charts that incorporate an early warning score system and its modifications to detect deterioration in adult patients in general medical and surgical wards. This may include extrapolation of data for general wards from studies conducted in multiple specialty sites on the exclusion list. • Educational/simulation settings • Admission may be via acute medical admission units/accident and emergency departments • 1 May 2007 to 23 May 2015. The start date was determined by the end date of the last systematic review by Odell (2009). • All research designs • English language • Studies evaluating the effects of early warning scoring systems on patient outcomes • Performance of EWS/MEWS • Reviews examining early warning scoring systems • All interventions and outcomes (except those listed in exclusions criteria). • Use of reporting guidelines or supplementary reporting guidelines • Effects of the EWS/MEWS system on rapid response systems (RRS)	• Studies conducted exclusively in: o Acute hospital wards o Intensive care units o High dependency units o Telemetry units o Accident and emergency departments o Specialist surgical wards o Psychiatric wards o Obstetric areas and wards o Paediatric areas and wards o Pre-hospital environments. • Studies only evaluating the effects of medical emergency teams on patient outcomes. • Observation charts in use that do no incorporate a MEWS or a reference standard i.e. not retrospective interpretation of data using a MEWS • Model development for predicting deterioration not incorporating EWS • Editorials, commentaries, symposium proceedings and non-peer reviewed journals. • Non-nurses using EWS

Additional search

The “related articles” option on PubMed was searched (albeit in a limited fashion)for additional relevant studies. A further Cochrane search was conducted in the Database of Systematic Reviews (Issue 1 of 12, January 2016), but not included in Table 2, using the indexing terms: “critical appraisal” and the Database of Abstracts of Reviews of Effects (DARE) for the terms “adequacy and reporting”.

Selection of relevant articles

Titles and abstracts of articles from keyword searches were screened for relevance. If full texts of relevant articles were available, in English and published between 1 May 2007 and 23 May 2015, these were included. Exclusion criteria were applied as listed in Table 2 and these refer to studies about EWS but not about vital signs, EWS/MEWS used outside adult general wards, reports, editorials or protocols on EWS/MEWS and if a study reported the use of EWS/MEWS by non-nurses.

Data extraction

In this section three steps of the data extraction process are explained and the results for each are presented in Tables 3, 4 and 5 respectively.

Grouping studies by research design, use of reporting guidelines and assessment of adequacy of reporting using a specific reporting guideline

A simple table was generated for grouping all eligible studies by methodological similarities, for assessing the use of a reporting guideline for studies on EWS/MEWS systems (yes/no; if yes, the name) and date of publication. Where the study design was not reported this was interpreted from the data.

Methodological characteristics of the studies reviewed

This included reporting the search strategy, use of experts, review method employed, how relevant references were selected, a flowchart of the selection of articles, results for each search, consensus between assessors, interrater agreement, meta-analysis performed, inclusion and exclusion criteria and reasons for exclusion of references.

General characteristics of the selected studies (if not included in reporting guideline and if no reporting guideline used)

General characteristics of the selected studies were used for evaluating the overall quality of the study reporting and this was particularly useful if reporting guidelines were not used. This included reporting the type of journal in which the study was published (Nursing or Medical/General/Other), whether statisticians were listed among the authors or if statistical assistance was reported, funding source, type of clinical area reported and physiological parameters that were measured with the EWS/MEWS.

Statistical analysis

The analysis was descriptive. Data were summarized as frequency and proportions (%).

RESULTS

Search strategy

A flowchart of the selected articles is shown in Figure 1.

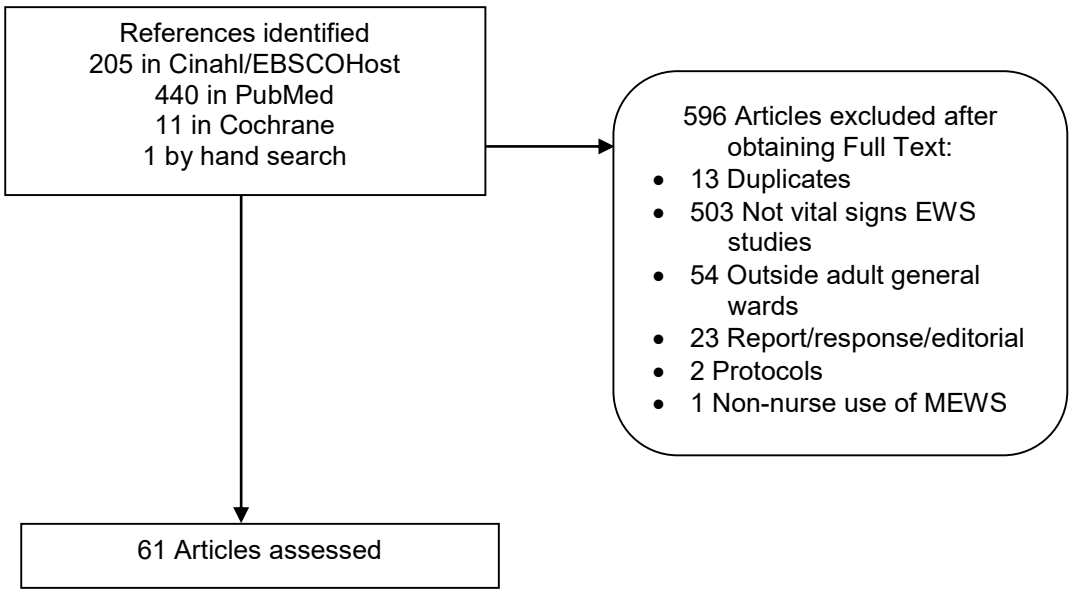

Figure 1: Flowchart of the selected articles

Of the 657 references identified from search terms, 596 articles were excluded after applying exclusion criteria, leaving 61 articles for assessment. Of the excluded articles the majority (n=503, 84.4%) were not studies about the use of EWS/MEWS for clinical deterioration in general wards. The search strategy is shown in Supplementary Information SI.

The additional search conducted in the Cochrane Database of Systematic Reviews using the indexing terms “critical appraisal” identified 18 references, and a search in DARE for the terms “adequacy and reporting” identified 79 references in Methods Studies and 212 in Trials. All references were excluded as not relevant to the seven key words.

Use of reporting guidelines for assessment of adequacy of reporting

Data in Table 3 show the study design (where reported or if not reported this was deduced from the methodology) (Grimes and Schulz, 2002), date of publication and use of a reporting guideline (yes/no) for studies on EWS/MEWS systems. The study designs listed in Table 3 are arranged according to reviews, followed by clinical studies organized according to two general categories: experimental and observational (Grimes and Schulz, 2002).

Table 3: Design of study and use of a reporting guideline

Design of study	Reporting guideline used Yes (%) N=61	Name of reporting guideline/Quality Assessment Tool used	During 2007-2010 n (%) N=61	During 2011-2015 n (%) N=61	Reporting guideline not used No (%) N=61
REVIEWS					
Systematic reviews	4 (6.6)	PRISMA SIGN QUADAS MQiPS and AHRQ	0	4 (6.6)	0
Integrative review	0		0	1 (1.6)	1 (1.6)
Narrative review	0		0	1	1 (1.6
Sub-total	4 (6.6)		0	6 (9.8)	2 (3.3)
EXPERIMENTAL STUDIES					
Pragmatic cluster randomised controlled trial	1 (1.6)	CONSORT	0	1 (1.6)	
Factorial design experiment	0		0	2* (3.3)	2 (3.3)
Prospective quasi-experimental trial	0		0	2 (3.3)	2 (3.3)
Pre- and post-intervention	0		0	2 (3.3)	2 (3.3)
Before-and-after controlled trial	0		0	1 (1.6)	1 (1.6)
Intervention time-series	0		0	1 (1.6)	1 (1.6)
Sub-total	1 (1.6)		0	9 (14.8)	8 (13.1)
OBSERVATIONAL STUDIES					
Various designs	1 (1.6)	STROBE	7 (11.5)	37 (60.7)	43 (70.5)
Mixed methods	0		0	2 (3.3)	2 (3.3)
Sub-total	1 (1.6)		7 (11.5)	39 (63.9)	45 (73.8)
TOTAL	6 (9.8)		7 (11.5)	54 (88.5)	55 (90.2)

Note on table:
PRISMA = Preferred Reporting Items for Systematic reviews and Meta-Analyses
SIGN = Scottish Intercollegiate Guidelines Network grading system
QUADAS = Quality Assessment of Diagnostic Accuracy Studies
MQiPS = Modified Quality in Prognosis Studies assessment tool; AHRQ = Agency for Healthcare Research and Quality
CONSORT = Consolidated Standards of Reporting Trials
STROBE = Strengthening the Reporting of Observational Studies in Epidemiology

Six of the 61 (9.8%) studies were reviews. The remaining 55 clinical studies on the use of EWS/MEWS systems for detecting deterioration in adult patients in general medical and surgical wards were mostly observational (46/55, 83.6%) rather than experimental (9/55, 16.4%). Most of the studies (54/61, 88.5%) were published after 2010 and all the studies using reporting guidelines were published after 2010.

Four (6.6%) reporting guidelines listed on the EQUATOR Network website were used in the reviewed studies on EWS/MEWS systems: PRISMA and QUADAS for systematic reviews, CONSORT for a RCT and STROBE for an observational study. One systematic review reported use of the Scottish Intercollegiate Guidelines Network (SIGN) grading system. Another systematic review reported use of the Modified Quality in Prognosis Studies (MQiPS) assessment tool and selected guidelines developed by the Agency for Healthcare Research and Quality. Flow diagrams for study recruitment and analysis were used in four observational studies, three between 2011 and 2015.

Methodological characteristics of the studies

None of the clinical studies reported a search strategy for the reviewed literature. Data presented in Table 4 therefore refer only to the search strategies for reviews (n=6).

Table 4: Methodological characteristics of reviews: search strategy

Methodology of the study	Overall, n (%) n=6
Search strategy reported	6 (100)
Electronic search	
MEDLINE	4 (66.7)
EMBASE	3 (50.0)
EBSCOHost: Cinahl	5 (83.3)
PubMed	2 (33.3)
The Cochrane Library/Central Register of Controlled Trials databases	4 (66.7)
Medion	1 (16.7)
Google	1 (16.7)
Hand search	1 (16.7)
Search of bibliography	2 (33.3)
References of systematic reviews	4 (66.7)
Experts	3 (50.0)
Review method	2 (33.3)*
Reporting of how relevant references were selected	4 (66.7)
Two or more persons independently	4 (66.7)
One person	0
One person with a quality assurance control	0
Two or more persons not independently	0
One person with a second person if difficulty experienced	0
Flowchart of the selection of articles	3 (50.0)
Results for each search	5 (83.3)
Reporting of consensus between assessors	3 (50.0)
Reporting of interrater agreement	1 (16.7)
Meta-analysis performed	0
Inclusion criteria reported	6 (100)
Exclusion criteria reported	5 (83.3)
Reasons for exclusion	4 (66.7)

(Adapted from Deschartes et al., 2011)

Note on table: * The PICO strategy was used to guide one search.

Electronic searches included mostly EBSCOHost: Cinahl (5/6, 83.3%), MEDLINE and The Cochrane Library/Central Register of Controlled Trials databases (4/6, 66.7%) respectively. Four of the six reviews (66.7%) reported references for systematic reviews, how relevant references were selected and that two or more persons independently selected the references. Three reviews (50.0%) reported the use of experts.

A flowchart of the selection of articles was reported in three of the six reviews (50.0%) and results for each search were reported in four of the reviews (66.7%). Consensus between assessors was reported in three of the six reviews (50.0%) and one review (16.7%) reported interrater agreement.

No meta-analyses were reported in the reviews. Inclusion criteria were reported in all six reviews and exclusion criteria in five reviews (83.3%) while reasons for exclusion of references were reported in four (80.0%) reviews.

General characteristics of the selected studies

Data in Table 5 show the general characteristics of the selected studies for evaluating the overall quality of study reporting and this was particularly useful if reporting guidelines were not used.

Table 5: General characteristics of selected studies

General characteristics of selected studies	Overall, n (%) N=61
Type of journal in which the study was published	
Nursing	14 (23.0%)
Medical/General/Other	47 (77.0)
Statisticians among the authors/statistical assistance	18 (29.5)
Funding source reported	26 (42.6)
No specific funding	33 (54.0)
Type of clinical area reported	
General wards	21 (34.4)
Wards: general/med + surg and specialities	7 (11.5)
Acute assessment unit	3 (4.9)
Acute wards outside ICU	1 (1.6)
Medical assessment units	2 (3.3)
Medical wards	8 (13.1)
Surgical wards	3 (4.9)
Medical + surgical wards	8 (13.1)
General wards/ICU	1 (1.6)
General wards/ICU/HDU	1 (1.6)
General wards/ICU/HDU/ED	1 (1.6)
Hospital training room/Non-clinical setting	2 (3.3)
Physiological parameters measured with EWS/MEWS	
None mentioned	2 (3.3)
2: HR, SBP	2 (3.3)
2: RR, LOC	1 (1.6)
2: RR, HR	1 (1.6)
2: sats + suppl 0_2	1 (1.6)
4: RR, Temperature, HR, LOC	1 (1.6)
4: Temperature, HR, SBP, RR	2 (3.3)
4: sats, HR, SBP, RR	1 (1.6)
5: RR, HR, BP, LOC, UO	2 (3.3)
5: RR, sats, HR, BP, Temp	4 (6.6)
5: RR, HR, BP, Temperature, LOC	1 (1.6)
5: Temperature, HR, SBP, RR, LOC	1 (1.6)
6: RR, sats, HR, BP, Temp, LOC	11 (18.0)
6: RR, HR, BP, Temperature, LOC, UO	2 (3.3)
6: RR, sats, suppl 0_2, HR, BP, LOC	4 (6.6)
6: BP,Temperature, HR, RR, sats, suppl 0_2	4 (6.6)
7: RR, sats, HR, BP, Temperature, LOC, UO	7 (11.5)
7: RR, sats, suppl O_2, HR, BP, Temperature, LOC	12 (19.7)
7: RR, sats, suppl O_2, HR, BP, UO, LOC	1 (1.6)
8: RR, sats, suppl O_2, HR, BP, Temperature, LOC, UO	3 (4.9)
Threatened airway	1 (1.6)
Worried about patient's condition	4 (6.6)

(Adapted from Dechartres et al., 2011)
Note on table: med + surg = medical and surgical, ICU = Intensive Care Unit, HDU = High Dependency Unit, ED = Emergency Department, HR = heart rate, SBP = systolic blood pressure, RR = respiratory rate, LOC = level of consciousness, sats = oxygen saturation, suppl O_2 = supplemental oxygen, UO = urine output, BP = blood pressure

Most EWS/MEWS and vital signs studies were published in non-nursing journals (47/61, 77.0%). Eighteen (29.5%) studies reported statistical assistance for data analysis. A funding source was not reported in the majority of studies (33/61, 54.0%). Most studies (21/61, 34.4%) were conducted in the general wards. A range of physiological parameters were reported but most of the studies mentioned respiratory rate, oxygen saturation, supplemental oxygen, heart rate, systolic blood pressure, temperature and level of consciousness (12/61, 19.7%). Four studies (6.6%) reported being worried about the patient's condition.

DISCUSSION

A critical appraisal is intended to assist reviewers and readers to assess the adequacy of published study methods (Altman, 2013). A review of the literature resulted in no evidence of a critical appraisal of the reporting of published clinical studies on EWS/MEWS systems for detecting deterioration in adult patients in the general medical and surgical wards. Checklists for critical appraisals are available (Altman, 2013, Dissemination, January 2009) but for the purpose of this critical appraisal, criteria used for assessing the quality of reporting in reviews of RCTs (Dechartres et al., 2011) seemed the most comprehensive and were adapted. Results show that most EWS/MEWS studies conducted in general wards were published after 2010 and in the category of observational designs. Some authors did not report the study type and this had to be deduced from the methodology (Grimes and Schulz, 2002). Few reporting guidelines were used in the published literature on EWS/MEWS systems.

Reporting guidelines are intended to help authors prepare better manuscripts and to assist peer reviewers to assess them (Hirst and Altman, 2012), particularly as it is reported that much published health care research is not useful, may be misleading, wasteful and even harmful (Simera et al., 2010, Chalmers and Glasziou, 2009). The studies on the use of reporting guidelines for the current critical appraisal were all published after 2010. Systematic reviews are guided by a strict scientific design to identify, evaluate and summarise the findings of relevant individual studies to provide more reliable estimates about the effects of interventions, thereby limiting the bias, methodologically flawed and context dependency associated with individual studies (Dissemination, January 2009). Two of four systematic reviews on EWS/MEWS systems used reporting guidelines listed on the EQUATOR Network: PRISMA (Alam et al., 2014) and QUADAS (Storm-Versloot et al., 2014). One systematic review (McNeill and Bryden, 2013) reported use of the Scottish Intercollegiate Guidelines Network (SIGN) grading system. In another systematic review authors (Smith et al., 2014) had adapted criteria in the Modified Quality in Prognosis Studies assessment tool for studies addressing predictive ability. For trials addressing health outcomes, components of the methods guide developed by the Agency for Healthcare Research and Quality were used. Neither the integrative review (Mapp et al., 2013) nor the narrative review (Kyriacos et al., 2011) alluded to the use of a reporting guideline.

Most reviewed clinical studies on EWS/MEWS systems did not use reporting guidelines. The only RCT of a MEWS system (Kyriacos et al., 2015) located in the published literature used the CONSORT (Campbell et al., 2004) reporting guideline. One observational study (Kyriacos et al., 2014) used the STROBE reporting guidelines. Flow diagrams of study recruitment and analysis were seldom used, making interpretation of the data difficult for replication and its usefulness for further research and practice questionable. Journals and publishers are therefore encouraged to increase awareness and to utilise reporting guidelines in their peer review process (Hirst and Altman, 2012).

Most of the reviewed studies were published in non-nursing journals, yet the monitoring, recording and interpretation of patients' vital signs is essentially a nursing responsibility (Hogan, 2006, Kisiel and Perkins, 2006) particularly as acutely ill patients are increasingly being nursed in general wards. An improvement in the quality of reporting of EWS studies should contribute towards improved evidence for improved vital signs monitoring practice and robust critical appraisal (Riegelman, 2013) of EWS studies. Improved practice includes an improvement in the nurses' clinical reasoning (Tanner, 2006) so that responses to clinical deterioration are not inappropriate (NHS NPSA, 2007) or delayed (Calzavacca et al., 2010, Calzavacca et al., 2008).

All the reviews reported a search strategy, but none of the clinical studies did so. Electronic searches included mostly CINAHL, MEDLINE and The Cochrane Library/Central Register of Controlled Trials databases respectively. This is not surprising as most nursing publications are located in the CINAHL database. It is encouraging that more than half the reviews reported references for systematic reviews, how relevant references were selected, results for each search and that two or more persons independently selected the references. Half the number of reviews reported the use of experts for the search strategy, a flowchart of the selection of articles and consensus between assessors. Experts are valuable as they may also be able to supply information about unpublished or ongoing research (Dissemination, January 2009). Interrater agreement was poorly reported in the reviews. Inclusion and exclusion criteria as well as reasons for exclusion of references were well reported in the reviews. A meta-analysis combines the results of individual studies thereby increasing power and precision in estimating intervention effects (Dissemination, January 2009). No meta-analyses were reported in the reviews, justifying concerns about the suboptimal reporting of published studies in general and of meta-analyses of RCTs in particular (Moher et al., 2009). A funding source was not reported in the majority of the studies. The influence of funding on reporting of research results is beyond the scope of this paper.

Most studies were conducted in general wards. A range of physiological parameters were reported but most of the studies reported respiratory rate, oxygen saturation, supplemental oxygen, heart rate, systolic blood pressure, temperature and level of consciousness. Clinical signs of deterioration (pallor, sweating, looking unwell) are very important as the MEWS does not replace the nurses' clinical judgement. Clinical signs of deterioration are incorporated into some MEWS charts although not scored (Kyriacos et al., 2011) so it was disappointing to find that only four studies reported being worried about the patient's condition.

LIMITATIONS

At review level limitations include restricting the search to English language papers (McNeill and Bryden, 2013). Increasingly, EWS/MEWS systems are being implemented in developing countries

(Rylance et al., 2009, Kyriacos et al., 2014, Zhu et al., 2015) and the non-English speaking world (Walcher et al., 2012) thereby underrepresenting the published literature from these countries in the review. The majority of the excluded articles were not studies about the use of EWS/MEWS system for clinical deterioration in general wards and a finer search strategy may have eliminated non-relevant references.

While there are published reviews undertaken by one reviewer (Monaghan, 2015) a more robust review includes two or more persons independently reviewing the references (Alam et al., 2014, McNeill and Bryden, 2013). Such persons should have a range of skills so that measures to minimize bias and error can be implemented at all stages of the review (Dissemination, January 2009).

CONCLUSIONS

Reporting of published studies on EWS systems reviewed for this critical appraisal, with the exception of reviews, was inadequate as most did not use reporting guidelines, thus limiting the use of study findings for developing clinical guidelines for teaching and in further research. Results suggest that future research on EWS systems in general hospital wards should be led and undertaken by nurses and should focus on experimental and intervention designs. In addition, nurse researchers, publishers and journals should be encouraged to use reporting guidelines to improve the quality of reporting studies. More robust reporting of EWS studies should improve the number and quality of reviews of such studies.

Appendix 1: Search strategy December 22, 2015.

	Search terms	Result
CINAHL	vital signs OR observations AND early warning score AND modified early warning score AND deterioration AND adult patients AND general wards	205
PUBMED	deterioration AND vital signs OR early warning score Filters: Publication date from 2007/05/01 to 2015/05/23; English; Adult: 19+ years	462
PUBMED	deterioration[All Fields] AND ("vital signs"[MeSH Terms] OR ("vital"[All Fields] AND "signs"[All Fields]) OR "vital signs"[All Fields]) OR (early[All Fields] AND warning[All Fields] AND score[All Fields]) AND ("loattrfull text"[sb] AND ("2007/05/01"[PDAT] : "2015/05/23"[PDAT]) AND English[lang] AND "adult"[MeSH Terms])	440
PUBMED	deterioration[All Fields] AND ("vital signs"[MeSH Terms] OR ("vital"[All Fields] AND "signs"[All Fields]) OR "vital signs"[All Fields]) OR (early[All Fields] AND warning[All Fields] AND score[All Fields]) AND (("loattrfull text"[sb] AND "loattrfree full text"[sb]) AND ("2007/05/01"[PDAT] : "2015/05/23"[PDAT]) AND English[lang] AND "adult"[MeSH Terms])	132

References

ALAM, N., HOBBELINK, E. L., VAN TIENHOVEN, A. J., VAN DE VEN, P. M., JANSMA, E. P. & NANAYAKKARA, P. W. 2014. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. *Resuscitation,* 85**,** 587-94.

ALTMAN, D. 2013. Critical appraisal checklists and reporting guidelines *The EQUATOR Network workshop.* WHO, Geneva

CALZAVACCA, P., LICARI, E., A., T. & ., E. A. 2010. The impact of Rapid Response System on delayed emergency team activation patient characteristics and outcomes - a follow-up study *Resuscitation* 81 31 – 35.

CALZAVACCA, P., LICARI, E., TEE, A. & AL, E. 2008. A prospective study of factors influencing the outcome of patients after a Medical Emergency Team review. *Intensive Care Med,* 34**,** 2112 – 2116.

CAMPBELL, M., ELBOURNE, D. & ALTMAN, D. 2004. CONSORT statement: extension to cluster randomised trials. *British Medical Journal,* 328**,** 702 - 8.

CHALMERS, I. & GLASZIOU, P. 2009. Avoidable waste in the production and reporting of research evidence. *Lancet,* 374**,** 86–89.

CIOFFI, J. 2000a. Recognition of patients who require emergency assistance: a descriptive study. *Heart & Lung : the Journal of Critical Care,* 29**,** 262-268.

CIOFFI, J. 2000b. Nurses' experiences of making decisions to call emergency assistance to their patients. *Journal of Advanced Nursing,* 32**,** 108-114.

DECHARTRES, A., CHARLES, P., HOPEWELL, S., RAVAUD, P. & ALTMAN, D. G. 2011. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol,* 64**,** 136-44.

DISSEMINATION, C. F. R. A. January 2009. Systematic Reviews. CRD's guidance for undertaking reviews in health care.

GRIMES, D. A. & SCHULZ, K. F. 2002. An overview of clinical research: the lay of the land. *The Lancet,* 359**,** 57-61.

HIRST, A. & ALTMAN, D. G. 2012. Are Peer Reviewers Encouraged to Use Reporting Guidelines? A Survey of 116 Health Research Journals. *PloS ONE,* 7**,** e35621.

HOGAN, J. 2006. Why don't nurses monitor the respiratory rates of patients? *British Journal of Nursing,,* 15**,** 489-492.

JOHNSTONE, C. C., RATTRAY, J. & MYERS, L. 2007. Physiological risk factors, early warning scoring systems and organizational changes. *Nursing in Critical Care,* 12**,** 219-224.

KISIEL, M. & PERKINS, C. 2006. Nursing observations: knowledge to help prevent critical illness. *British Journal of Nursing,,* 15**,** 1052-1056.

KYRIACOS, U., JELSMA, J., JAMES, M. & JORDAN, S. 2014. Monitoring vital signs: development of a modified early warning scoring (MEWS) system for general wards in a developing country. *PLoS One,* 9**,** e87073.

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,* 16**,** 103.

KYRIACOS, U., JELSMA, J. & JORDAN, S. 2011. Monitoring vital signs using early warning scoring systems: a review of the literature. *Journal of Nursing Management,* 19**,** 311-330.

LIBERATI, A., ALTMAN, D. G., TETZLAFF, J., MULROW, C., GOTZSCHE, P. C., IOANNIDIS, J. P. A., CLARKE, M., DEVEREAUX, P. J., KLEIJNEN, J. & MOHER, D. 2009. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *Annals of Internal Medicine,* 151**,** W-65-W-94.

MAPP, I. D., DAVIS, L. L. & KROWCHUK, H. 2013. Prevention of Unplanned Intensive Care Unit Admissions and Hospital Mortality by Early Warning Systems. *Dimensions of Critical Care Nursing,* 32**,** 300-309.

MCNEILL, G. & BRYDEN, D. 2013. Do either early warning systems or emergency response teams improve hospital patient survival? A systematic review. *Resuscitation,* 84**,** 1652-67.

MELLO, M. M., STUDDERT, D. M. & BRENNAN, T. A. 2003. The New Medical Malpractice Crisis. *New England Journal of Medicine,* 348**,** 2281-2284.

MOHER, D., COOK, D. J., EASTWOOD, S., OLKIN, I., RENNIE, D. & STROUP, D. F. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet,* 354**,** 1896-1900.

MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. & GROUP, F. T. P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal,* 339**,** b2535.

MONAGHAN, T. 2015. A critical analysis of the literature and theoretical perspectives on theory–practice gap amongst newly qualified nurses within the United Kingdom. *Nurse Education Today,* 35**,** e1-e7.

NHS NPSA. 2007. *Safer Care for the Acutely Ill Patient* [Online]. Available: http://www.npsa.nhs.uk/patientsafety/alerts-and-directives/directives-guidance/acutely-ill-patient/ [Accessed 2008/07/10

ODELL, M., VICTOR, C. & OLIVER, D. 2009. Nurses' role in detecting deterioration in ward patients: systematic literature review. *Journal of Advanced Nursing,* 65**,** 1992-2006.

REASON, J. 2000. Human error: models and management. *British Medical Journal.,* 320**,** 768.

RIEGELMAN, R. K. C. A., RINKE, M.L. (2013). STUD 2013. *Studying a study & testing a test: Reading evidence-based health research* Riegelman, R.K. Contributing author, Rinke, M.L. (2013). Studying a study & testing a test: Reading evidence-based health research (). Philadelphia: Wolvers Kluwer. Available: http://www.rhc.ac.ir/Files/Download/pdf/nursingbooks/Studying%20A%20Study%20and%20Testing%20a%20Test%20Reading%20Evidence-based%20Health%20Research-2013-CD.pdf [2015, 12 December], Philadelphia: Wolvers Kluwer.

ROYAL COLLEGE OF PHYSICIANS 2012. National Early Warning Score (NEWS): standardising the assessment of acute-illness severity in the NHS. *Report of a working party.* London:RCP: Royal College of Physicians.

RYLANCE, J., BAKER, T., MUSHI, E. & MASHAGA, D. 2009. Use of an early warning score and ability to walk predicts mortality in medical patients admitted to hospitals in Tanzania. *Trans R Soc Trop Med Hyg,* 103**,** 790-4.

SIMERA, I., MOHER, D., HIRST, A., HOEY, J., SCHULZ, K. & AL., E. 2010. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med,* 8**,** 24.

SMITH, M. E., CHIOVARO, J. C., O'NEIL, M., KANSAGARA, D., QUINONES, A. R., FREEMAN, M., MOTU'APUAKA, M. L. & SLATORE, C. G. 2014. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. *Ann Am Thorac Soc,* 11**,** 1454-65.

STORM-VERSLOOT, M. N., VERWEIJ, L., LUCAS, C., LUDIKHUIZE, J., GOSLINGS, J. C., LEGEMATE, D. A. & VERMEULEN, H. 2014. Clinical Relevance of Routinely Measured Vital Signs in Hospitalized Patients: A Systematic Review. *Journal of Nursing Scholarship,* 46**,** 39-49 11p.

SUBBE, C., DAVIES, R., WILLIAMS, E., RUTHERFORD, P. & GEMMELL, L. 2003. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. *Anaesthesia,* 58**,** 797 - 802.

TANNER, C. A. 2006. Thinking Like a Nurse: A Research-Based Model of Clinical Judgment in Nursing. *Journal of Nursing Education,* 45**,** 204-211.

WALCHER, F., KULLA, M., KLINGER, S., ROHRIG, R., WYEN, H., BERNHARD, M., GRAFF, I., NIENABER, U., PETERSEN, P., HIMMELREICH, H., SCHWEIGKOFLER, U., MARZI, I. & LEFERING, R. 2012. [Standardized documentation in emergency departments with the core dataset of the DIVI]. *Unfallchirurg,* 115**,** 457-63.

WILSON, R. M., HARRISON, B. T., GIBBERD, R. W. & HAMILTON, J. D. 1999. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *The Medical Journal of Australia,* 170**,** 411-415.

ZHU, X., LI, T. & LIU, M. 2015. [Evaluation and analysis of monitoring and early warning functions of the occupational disease reporting system in China]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi,* 33**,** 422-6.

ZIMLICHMAN, E., SZYPER-KRAVITZ, M., UNTERMAN, A., GOLDMAN, A., LEVKOVICH, S. & SHOENFELD, Y. 2009. How is My Patient Doing? Evaluating Hospitalized Patients Using Continuous Vital Signs Monitoring. *Israeli Medical Association Journal,* 11**,** 382-384.