The impact of face-to-face and web-based simulation on patient deterioration and patient safety.

Funded by the Victorian Government

Project Team

Project Leads:	Professor Simon Cooper (Federation University) Professor Leigh Kinsman (University of Tasmania and Tasmanian Health Service)
Project Manager:	Ms Catherine Chung
Research Fellow:	Dr Robyn Cant
Team Members:	Ms Jayne Boyle Ms Amanda Cameron A/Professor Penny Cash Mr Cliff Connell Ms Lisa Evans Dr Jeong Ah Kim Ms Angela McKay Ms Denise McInnes Ms Lisa Norman Dr Erica Penz Dr Thomas Rotter

ISBN: **978-1-876851-972 (eBook)**

Executive Summary

There are significant patient safety concerns regarding nurses' ability to identify and manage patients that are physiologically deteriorating. Recent evidence from the 'failure to rescue' literature indicates a high level of disturbed physiological variables in the general ward population and poor patient outcomes due to mismanaged patient deterioration. However, it is known that educational outcomes can be improved with approaches that emphasise active learning approaches in simulated and clinical environments. In relation to this, team members from this project completed a systematic literature review identifying that blended learning approaches, with medium to high fidelity simulation, had a positive educational impact. It is apparent, however, that less is known about the impact of web based programs, the clinical impact of simulation programs and the cost benefit of educational approaches.

In this mixed methods interventional cohort trial the aim was to compare the effectiveness of two forms of simulation education, face-to-face (F2F) versus web-based (WB), on nurses' ability to detect and manage patient deterioration.

Australian nurses, from the State of Victoria, were trained (for up to two hours) in primary responses to emergencies in medical wards in two regional public hospitals and two private hospitals using either F2F or a WB version of the patient deterioration program FIRST^{2}ACT™ (Feedback Incorporating Review and Simulation Techniques to Act on Clinical Trends). The F2F program was completed with teams of three nurses, while the WB program was completed individually.

Incorporating independent performance ratings, patient notes review and focus group interviews we measured educational impact, clinical impact (e.g. compliance with vital signs chart recording and clinical review criteria) and economic impacts.

Results

A total of 129 nurses attended training (55 WB; 74 F2F) from a population of 141 (91% attendance rate)

Education Impact

Web based participant outcomes

- Knowledge, skills, confidence and competence increased significantly ($p=<0.001$) with a median program evaluation score of 4/5

Face-to-face participant outcomes

- Knowledge, confidence and competence increased significantly ($p=<0.001$) with a median program evaluation score of 5/5 (a higher rating than the WB program)
- Skill performance did not improve however the teams of three only led one of the three scenarios once.

Clinical Impact

A total of 1,951 medical notes and vital sign charts were reviewed at the four hospital sites – 19.5% were excluded (e.g. patients 'not for resuscitation') leaving 1,564 patient records for analysis. Fifty percent - 50.1% (n= 783) were admissions prior to the teaching interventions

and 49.9% (n= 781) were after, with no key difference in patient characteristics pre – post intervention (e.g. age, sex).

Around 60% reached clinical review criteria (the requirement that senior staff are informed) pre and post intervention *[Identified as a coloured area on the vital signs charts e.g. a low BP or fast heart rate].*

Of those that met clinical review criteria:

- 12% (n=56) were correctly reviewed pre-intervention
- 27% (n=130) were correctly reviewed post intervention (p=<0.001). (i.e. a doubling + of the review rate). This was confirmed in time series analyses (TSA) which demonstrated a significant interventional effect.
- There was no difference in 'effect' between the web based and face to face interventions (i.e. both were equally effective in increasing clinical reviews)

Of those patients meeting Medical Emergency Team (MET) calling criteria:

- There was no statistically significant difference in the rate of MET team activations between pre (18/53: 34%) and post audits (19/45: 42%). The number of patients reaching MET call criteria reduced by 15% from 53 to 45 following the simulation intervention.

Other impacts (both F2F and WB) that were identified through TSA:

- Inappropriate oxygen therapy (e.g. nasal prongs instead of a mask) declined pre-intervention and in the first two months' post interventions, but was not maintained beyond this point.
- An improving trend in the frequency of vital signs recording pre interventions was maintained following the interventions.
- A significant improvement in the recording of pain scores.
- More 'at-risk' patients were reviewed within 60 mins post intervention.
- More interventions were performed post intervention.

Economic Outcomes

Training costs:

- The WB version costs more to implement initially (including the cost of program development) but reduces significantly over time in comparison to F2F training. For example the cost of training 'the first' 300 - F2F participants is a mean of $77.16 compared to web based costs of $56.15

Hospital costs:

- There was no significant difference in the mean patient hospitalisation costs related to the two interventions. Patient length of stay was longer in the two wards where web based training was completed but this was unrelated to the training interventions.

Focus Group Findings Ten focus groups were completed at the four hospital sites (65 staff). Three main themes emerged:

- **'Program impact':** Both programs were positively evaluated but with the desire for improved debriefing and concerns over the fidelity of simulation.

- **‘Managing deterioration’:** Respondents raised concerns about clinical communication especially in relation to the poor recording of altered MET criteria.
- **‘Clinical judgement’.** Respondents identified that the use of the prescribed vital signs charts reduced individuals’ critical thinking flexibility/judgement.

Conclusions

- Both programs were positively evaluated by participants and significantly improve knowledge, confidence and competence.

- The WB program enhances skill development whilst F2F training is likely to improve skills with repetitive practice.

- Both programs had a significant clinical impact, increasing clinical reviews and improving nursing interventions.

- Over time the web based version will be considerably cheaper to implement.

- Interviews indicate positive educational and clinical impact but room for improved debriefing, limits to the fidelity of simulation, and a need to consider clinical communication, vital sign chart forms and practice.

The future

- To improve nurses’ skills in detecting and managing patient deterioration, the use of BOTH programs consecutively is recommended i.e. ‘blended learning’.

- Bearing in mind the continued poor compliance to vital signs recording/interventions there is a need to review charts and response criteria.

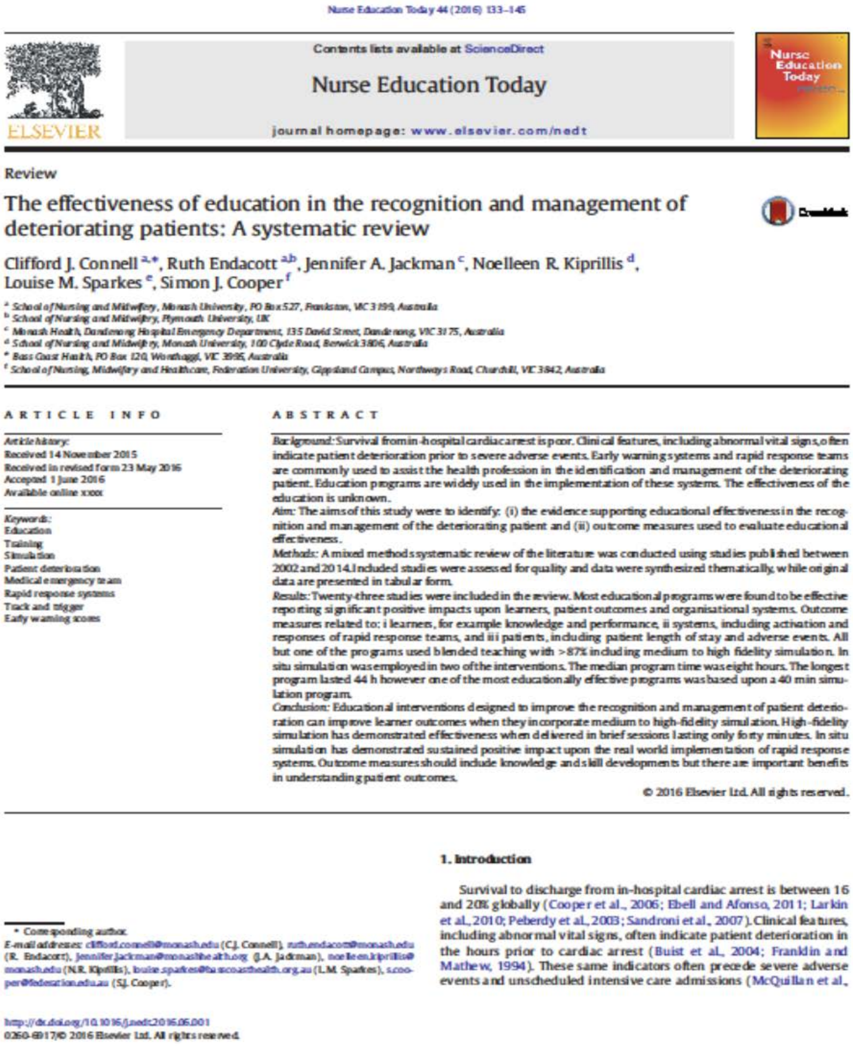
- Expansion of these simulation programs to inter-disciplinary teams would add to our understanding of how simulation can contribute to better patient outcomes in inter-disciplinary environments.

Introduction/Background

The Australian Commission on Safety and Quality in Health Care (ACSQH) has developed the National Safety and Quality Health Service Standards (Australian Commission on Safety and Quality in Healthcare 2012). These aim to protect the public from harm and to improve the quality of health service provision, with a specific focus on healthcare professionals' ability to recognise and respond to 'Clinical Deterioration in Acute Health Care' (Standard 9). In recent years the acuity of patients managed on general wards in acute facilities has increased whilst access to critical care beds has decreased (Needleman, 2013). There is a high level of disturbed physiological variables in the general ward population (Harrison et al, 2005) and poor patient outcomes due to mismanaged patient deterioration (Harrison et al, 2005; Endacott & Westley 2006; Hogan et al, 2006; Waldie et al, 2016). Missed indicators of deterioration have also been noted in patients following discharge from hospital and in smaller rural hospitals (Endacott & Westley, 2006) whilst further evidence has indicated that nurses are not always clear about when to call for assistance (Cioffi 2000), do not seek advice and fail to appreciate clinical urgency (Waldie et al, 2016). In addition, smaller outer urban and regional hospitals receive high acuity patients less frequently than their metropolitan counterparts with the risk of skill decline and skill mix concerns, whilst an aging population increases the likelihood of co-morbidities. These issues emphasize the need for good clinical assessment skills to identify deterioration earlier.

It is known, however, that the management of deteriorating patients can be influenced by education and past experience (Connell et al, 2016; Liaw et al, 2016; O'Leary et al, 2016). In health care contemporary education approaches emphasise the need for active learning (Endacott et al, 2003) and increased use of simulated environments (Motola & Devine 2013) to reduce medical errors (Ziv et al, 2005) in settings which have high environmental and psychological fidelity (believability) (Fritz & Flanagan, 2008). Further, in a 2016 systematic review (Connell et al, 2016) we identified 23 studies that reported on outcomes from education programs. Most were found to be effective and incorporated blended learning approaches (87%) with medium to high fidelity simulation (Figure 1).

Figure 1: Impact of Patient Deterioration Programs


Nurse Education Today 44 (2016) 133–145

Contents lists available at ScienceDirect

Nurse Education Today

journal homepage: www.elsevier.com/nedt

Review

The effectiveness of education in the recognition and management of deteriorating patients: A systematic review

Clifford J. Connell [a,*], Ruth Endacott [a,b], Jennifer A. Jackman [c], Noelleen R. Kiprillis [d], Louise M. Sparkes [e], Simon J. Cooper [f]

[a] School of Nursing and Midwifery, Monash University, PO Box 527, Frankston, VIC 3199, Australia
[b] School of Nursing and Midwifery, Plymouth University, UK
[c] Monash Health, Dandenong Hospital Emergency Department, 135 David Street, Dandenong, VIC 3175, Australia
[d] School of Nursing and Midwifery, Monash University, 100 Clyde Road, Berwick 3806, Australia
[e] Bass Coast Health, PO Box 120, Wonthaggi, VIC 3995, Australia
[f] School of Nursing, Midwifery and Healthcare, Federation University, Gippsland Campus, Northways Road, Churchill, VIC 3842, Australia

ARTICLE INFO

Article history:
Received 14 November 2015
Received in revised form 23 May 2016
Accepted 1 June 2016
Available online xxxx

Keywords:
Education
Training
Simulation
Patient deterioration
Medical emergency team
Rapid response systems
Track and trigger
Early warning scores

ABSTRACT

Background: Survival from in-hospital cardiac arrest is poor. Clinical features, including abnormal vital signs, often indicate patient deterioration prior to severe adverse events. Early warning systems and rapid response teams are commonly used to assist the health profession in the identification and management of the deteriorating patient. Education programs are widely used in the implementation of these systems. The effectiveness of the education is unknown.
Aim: The aims of this study were to identify: (i) the evidence supporting educational effectiveness in the recognition and management of the deteriorating patient and (ii) outcome measures used to evaluate educational effectiveness.
Methods: A mixed methods systematic review of the literature was conducted using studies published between 2002 and 2014. Included studies were assessed for quality and data were synthesized thematically, while original data are presented in tabular form.
Results: Twenty-three studies were included in the review. Most educational programs were found to be effective reporting significant positive impacts upon learners, patient outcomes and organisational systems. Outcome measures related to: i learners, for example knowledge and performance, ii systems, including activation and responses of rapid response teams, and iii patients, including patient length of stay and adverse events. All but one of the programs used blended teaching with >87% including medium to high fidelity simulation. In situ simulation was employed in two of the interventions. The median program time was eight hours. The longest program lasted 44 h however one of the most educationally effective programs was based upon a 40 min simulation program.
Conclusion: Educational interventions designed to improve the recognition and management of patient deterioration can improve learner outcomes when they incorporate medium to high-fidelity simulation. High-fidelity simulation has demonstrated effectiveness when delivered in brief sessions lasting only forty minutes. In situ simulation has demonstrated sustained positive impact upon the real world implementation of rapid response systems. Outcome measures should include knowledge and skill developments but there are important benefits in understanding patient outcomes.

1. Introduction

Survival to discharge from in-hospital cardiac arrest is between 16 and 20% globally (Cooper et al., 2006; Ebell and Afonso, 2011; Larkin et al., 2010; Peberdy et al., 2003; Sandroni et al., 2007). Clinical features, including abnormal vital signs, often indicate patient deterioration in the hours prior to cardiac arrest (Buist et al., 2004; Franklin and Mathew, 1994). These same indicators often precede severe adverse events and unscheduled intensive care admissions (McQuillan et al.,

* Corresponding author.
E-mail addresses: clifford.connell@monash.edu (C.J. Connell), ruth.endacott@monash.edu (R. Endacott), jennifer.jackman@monashhealth.org (J.A. Jackman), noelleen.kiprillis@monash.edu (N.R. Kiprillis), louise.sparkes@basscoasthealth.org.au (L.M. Sparkes), s.cooper@federation.edu.au (S.J. Cooper).

http://dx.doi.org/10.1016/j.nedt.2016.06.001

The researchers and educators in this study have worked on a range of patient safety initiatives covering non-technical skills and patient deterioration management with students, midwives and qualified rural nurses who completed face-to-face and web based training (Cooper et al, 2010; Endacott et al, 2010; Buykx et al, 2011; Kinsman et al, 2012; Scholes et al, 2012). In this work we examined patient deterioration management in a simulated setting using an interventional program known as First2Act(Web) (Feedback Incorporating Review and Simulation Techniques to Act on Clinical Trends) http://first2actweb.com/

The face-to-face program and web-based version includes a course manual, assessment tests, simulation with patient actors (simulated patients) and feedback techniques, delivered to up to three participants over a 1-2 hour period. Our findings indicate that participants have reasonable theoretical knowledge but that they fail to respond appropriately; for example as patients deteriorate and anxiety levels increase, performance considerably reduces. Our findings, however, also indicate that the educational experience, which includes knowledge and simulation based skill tests (Objective Structured Clinical Examinations -OSCEs), reflection and feedback, has a significant impact on participants' learning. Further, in a small face-to-face training study with registered nurses we included patient notes review and a time series analysis, identifying a notable improvement in patient management skills (clinical impact) with significant improvements in the appropriate frequency of vital signs recording (p=0.009), charting of pain scores (p=0.003), and oxygen delivery (Kinsman et al, 2012).

These findings indicated that educational outcomes and patient management can be significantly improved with applicable education interventions. However these studies were resource intensive as they were delivered face-to-face in clinical skills centres. We therefore developed an interactive web based e-simulation model (funded by the Australian Office for Teaching and Learning) which has now (Nov 2016) been completed by 10,000+ participants from across the world.

[see 'how the scenarios work' http://first2actweb.com/index.php/first2actweb/.]

In this current version three patient simulations are included covering myocardial infarction, COPD and hypovolemia. In Version 2 and 3 to be released in 2017, we include maternal, neurological and diabetes emergencies. All scenarios feature simulated patient actors, as opposed to manikins, which increases the fidelity of the scenarios.

Based on these programs our intention was to train Australian nurses from the State of Victoria, in primary (the first 8 minutes) responses to emergencies. The settings were a public medical ward at LaTrobe Regional Hospital (LRH), Central Gippsland Health Service [CGHS] – Sale, and two private wards at St John of God Hospitals (Berwick and Bendigo) using either in-situ face-to-face or web based version of First2Act. In actuality, we were unable to complete all the training 'in situ' (e.g. ward based) as there were rarely sufficient bed vacancies to enable access. We therefore used the respective hospitals' simulation centres which were equipped to mirror the hospital wards.

The impact of these interventions on patient outcomes, nursing practice and costs was determined through quantitative (medical records audit and economic analyses) and qualitative approaches (focus groups with participating nurses). The study did initially focus upon nurse education and impact but it is anticipated that the findings will inform the development and impact of training programs for other health professions.

In summary it is expected that this project will have a positive impact on patient safety, and improve quality and health care outcomes, and explore the relative impact and cost benefits of face to face and web based programs.

Design and Methods

A summary of the methods are described below and the study protocol has been published on line – open access- and is available from https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-016-1683-0

See Figure 2:

Figure 2: Study Protocol

Cooper et al. BMC Health Services Research (2016) 16:475
DOI 10.1186/s12913-016-1683-0

BMC Health Services Research

STUDY PROTOCOL Open Access

The impact of web-based and face-to-face simulation on patient deterioration and patient safety: protocol for a multi-site multi-method design

Simon J. Cooper[1*], Leigh Kinsman[2], Catherine Chung[1], Robyn Cant[1], Jayne Boyle[3], Loretta Bull[4], Amanda Cameron[5], Cliff Connell[6], Jeong-Ah Kim[1], Denise McInnes[7], Angela McKay[2], Katrina Nankervis[4], Erika Penz[8] and Thomas Rotter[9]

Abstract

Background: There are international concerns in relation to the management of patient deterioration which has led to a body of evidence known as the 'failure to rescue' literature. Nursing staff are known to miss cues of deterioration and often fail to call for assistance. Medical Emergency Teams (Rapid Response Teams) do improve the management of acutely deteriorating patients, but first responders need the requisite skills to impact on patient safety.

Methods/design: In this study we aim to address these issues in a mixed methods interventional trial with the objective of measuring and comparing the cost and clinical impact of face-to-face and web-based simulation programs on the management of patient deterioration and related patient outcomes. The education programs, known as 'FIRST[2]ACT', have been found to have an impact on education and will be tested in four hospitals in the State of Victoria, Australia. Nursing staff will be trained in primary (the first 8 min) responses to emergencies in two medical wards using a face-to-face approach and in two medical wards using a web-based version FIRST[2]ACTWeb. The impact of these interventions will be determined through quantitative and qualitative approaches, cost analyses and patient notes review (time series analyses) to measure quality of care and patient outcomes.

Discussion: In this 18 month study it is hypothesised that both simulation programs will improve the detection and management of deteriorating patients but that the web-based program will have lower total costs. The study will also add to our overall understanding of the utility of simulation approaches in the preparation of nurses working in hospital wards. (ACTRN12616000468426, retrospectively registered 8.4.2016).

Keywords: Clinical audit, Education, Nursing, E-simulation, Patient safety, Simulation-based learning

* Correspondence: s.cooper@federation.edu.au
[1]School of Nursing Midwifery and Healthcare, Federation University Australia, Churchill, Victoria 3842 and Mt. Helen, Victoria 3350, Australia
Full list of author information is available at the end of the article

Aim:

To compare the effectiveness of two forms of simulation education (face-to-face versus web-based) on nurses' ability to detect and manage patient deterioration.

[The reader should note that the programs, though similar, used a different form of delivery and were in themselves not directly comparable. As such the intention was to measure the impact of each program rather than to compare the programs per se.]

Hypotheses:

The research hypotheses to be tested are:

1. That both models of simulation will improve the detection and management of patient deterioration and patient outcomes.
2. That the web-based program will have superior cost-effectiveness in improving the detection and management of patient deterioration and patient outcomes.

Objectives:

1. To evaluate the educational outcomes of the face-to-face and web-based simulation programs.
2. To measure and compare the impact of the face-to-face and web-based simulation programs on the detection and management of patient deterioration, and patient outcomes.

3. To measure and compare the cost-effectiveness of the face-to-face and web-based simulation programs.

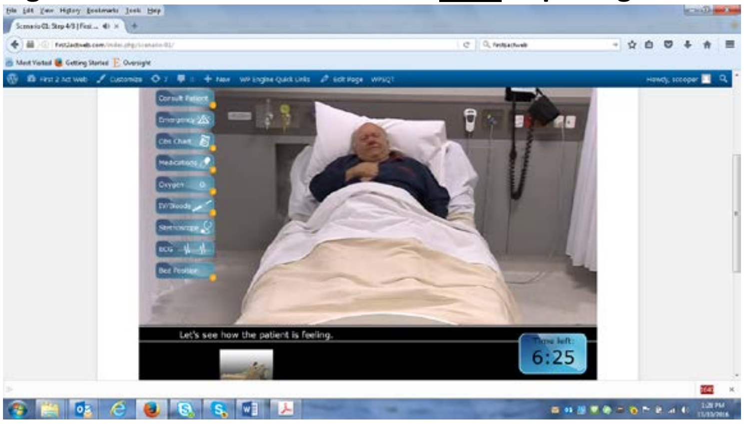
Design: a mixed methods interventional cohort trial comparing the impact of two educational interventions within and between groups.

The *primary outcomes* measured were the 1) proportion of patients reaching 'clinical review' criteria and 'Medical Emergency Team (MET) calling criteria and 2) the quality of nursing assessment. The secondary outcomes measured were 1) cost of provision of face-to-face versus web based programs and 2) stakeholders' views and attitudes to the forms of program delivery and impact.

The Interventions

The First2Act web-hosted 'e-simulation' (screen based) program includes a range of educational material, pre and post course assessments, and three interactive scenarios (cardiac, shock and respiratory cases) which run over eight minutes. Patient actors depict a deteriorating patient who significantly deteriorate at the four minute mark. Individual participants are require to 'click' on various actions – such as inserting an IV line, taking vital signs, recording an ECG or selecting a form of oxygen therapy – resulting in pop up videos of each action (Figure 3). Detailed feedback on performance is provided at the end of each scenario and at the end of the program, which takes approximately 1-1.5 hours to complete. Staff completed the program in their own time or were released from ward duties to do so.

Figure 3: Screen shot of First2Act<u>Web</u> depicting action buttons and patient actor



The First2Act face-to face version mirrors the content of the web based program. Set in a clinical lab, designed to mirror a ward environment, it includes the same three scenarios which are again portrayed by a simulated patient actor. However, in this version, we trained lecturing staff to act out each role and the same individual then rated participant performance after each scenario. Ratings were verified with a second lecturer who supported, but did not prompt, participants and undertook medical tasks as directed by the nurse participants (e.g. inserted an IV line or prescribed medication). Scenarios were filmed on a tablet and using these recordings a third lecturer spent approximately 30 minutes debriefing participants. In total the program took two hours to complete. In contrast to the web based program the face-to-face program included teams of three individuals with the leadership role rotated between the three scenarios/participants. This increased the viability of the training and enabled rating and development of teamwork. Table 1 lists the detailed stages of both programs. Staff were rostered onto the two hour program over a working day with their ward work covered by clinical educators or additional sessional staff.

Table 1: Program content for each intervention (Adapted from Cooper et al 2016)

	Group 1 **First^{2}Act Web**	**Group 2** **First^{2}Act Face-to-face**
Introductory preamble	Short written introduction with explanation.	Short verbal introduction with explanation.
Demographics and Pre-course MCQ	A demographics survey and a 12 item multiple choice questionnaire (MCQ).	As in web-based.
Background material	An animated slide show – with voice over from an educator. A variety of styles include pop up illustrations, video links, clear transitions and summaries.	As in web-based.
Scenarios (skills)	Three interactive web-based scenarios (8 minutes each) Each scenario filmed using patient actors trained to simulate symptoms of deterioration and incorporating applicable moulage (e.g. cyanosis). At the mid-point of each scenario the 'patient' declines significantly. The technology supports a variety of actions presented in real time, including pop up videos of vital sign recordings, drug administration, oxygen delivery, and bed position. Performance rated on line based on pre-determined criteria.	Three interactive scenarios (8 minutes each) conducted in a clinical skills centre using patient actors and video-recorded to enable participant review. As in web-based. Performance rated by two researchers at the end of each scenario – using predetermined criteria/check lists (an OSCE format).
Feedback	The software gathers all performance data for automatized feedback at the end of each scenario.	Face-to-face feedback conducted with an educator using video and performance records.
Post-course MCQ	Repeat of the MCQ.	As in web-based.

Course evaluation	A course evaluation and reflective review of educational impact	As in web-based.
Certification	Download a course participation certificate	Issued with a course participation certificate.
Course Manual	Download the course manual for reflection and review	Issued with the course manual for reflection and review
Time commitment	1 to 1.5 hours	2 hours

Setting: The medical wards of LRH, CGHS Sale and St John of God hospitals, Berwick and Bendigo. Two medical wards were assigned as the web based intervention Group 1 (Sale & Bendigo), while the face-to-face intervention (Group 2) was allocated to two medical wards at LRH & Berwick. By taking this approach we ensured that there is no contamination between homogeneous groups as staff did not work between hospitals.

Population and intervention sample size: All qualified nursing staff (Registered and Enrolled Nurses) working in general medical wards were eligible: LRH = 47; Sale = 32; SJOG Bendigo = 34; SJOG Berwick = 28. In pilot hospital-based studies we had achieved a participation rate of 82% and anticipated similar participation rates for this study.

Clinical impact measures and sample: As described above we measured the proportion of patients reaching 'clinical review' criteria and 'Medical Emergency Team (MET) calling criteria; the quality of nursing assessment; the cost of provision of face-to-face versus web based programs and stakeholders' views and attitudes to the forms of program delivery and impact.

As per our previous studies, and as per validated data extraction tools, we used appropriate frequency of vital signs and evidence-based responses to deterioration as proxies for quality of nursing care. In addition, we assessed the costs associated with the provision of face-to-face versus web based learning and investigated stakeholder views and attitudes through semi-structured interviews with stakeholder groups at each site.

Patient sample size: In line with time series analyses guidelines and initial modelling (Kinsman et al, 2012) we aimed to extract a random selection of 100 admissions (using electronic random numbers) from each health services' participating ward for each of three months prior to, and three months after the training. Across sites the sample ranged from approximately 64% to 89% of all admissions during each one-month period. The first 80 episodes, allowing for some redundant cases which did not meet the inclusion criteria, were included in the records audit.

Intervention measures

Quantitative data was captured from the following sources:

1. **Patient medical records/vital signs charts.**

We developed and tested a chart audit tool based on a rationale for item inclusion, definitions, and clinical expert ratings of relevance and clarity which achieved a high Content

Validity Index of 0.83. Details related to the development of this tool are available in the publication by McConnell-Henry, et al. (2015) In this latest study an expert panel of six researchers/clinicians updated the tool by altering several variables to be collected in line with survey objectives.

This tool was uploaded to Qualtrics (as an on line survey) to enable us to collect data relating to patient demographics, whether or not patients met clinical review and MET calling criteria, and what actions were performed in relation to these outcomes, for example team activations, vital signs recorded, oxygen given, interventions recorded.

2. A multiple choice knowledge questionnaire (MCQ).

A validated 12 item Multiple Choice Questionnaire (pre-post) (Cooper et al, 2010) completed by participants before and after the program in order to measure knowledge regarding patient deterioration. This measure enabled identification of pre-existing knowledge in relation to skill performance during the simulation exercises, and any change in knowledge based on program participation.

3. Skills assessments.

All participants completed three contrasting simulation exercises that were recorded as interactive web based versions or face-to-face with a patient actor. Each scenario depicted common presenting conditions (e.g. AMI, COPD, hypovolaemia). During the simulation exercises, information was presented in a manner that most clearly reflected the real world requiring participants to be an active searcher. This approach enhanced the ecological validity of the simulation (Cioffi 2001) allowing the participant to experience clinical thinking in a dynamic manner. Levels of relevant information and levels of uncertainty was taken into account and incorporated into the scenarios. Criteria for item inclusion were set by a panel of clinical experts in line with the design of Objective Structured Clinical Examinations (OSCE). Scoring for both programs used a nominal approach – performed yes/no – with the total score for each scenario summed.

4. Participant evaluations.

Each program was evaluated by participants on completion using a mix of rating scales and open ended comment options. Pre-post evaluations of confidence and competence were also sort.

5. Costs derived from patient records and inventory of resources.

We conducted a cost analysis (Ramsey et al, 2005; Glick et al, 2007) of web-based (WB) versus face-to-face (F2F) training in terms of mean difference in training costs and difference in hospitalization costs. Training intervention costs were captured for each group including program development, software development, and educator commitment and staff release costs. Using a chart audit pre and post training intervention of each participating hospital ward, length of stay, hospital costs (based on AR-DRG codes and resource intensity weights), and ICU admissions were captured.

Total and mean 6 month training costs are reported for each training group. Projected total and mean training costs for each group are modeled with 1) an equivalent number of trainees and 2) as additional nurses are trained.

Mean Length of stay (LOS), mean hospitalization costs and the difference in means are reported for the pre-intervention and post-intervention period within each group (WB and F2F). The *between group difference* in mean LOS and hospitalization cost is also reported using the difference-in-differences methodology (Appendix 2). 95% confidence intervals around the mean differences in LOS and costs between the two groups were performed using bootstrap sampling.

Qualitative data
Qualitative data about experience of the training intervention was captured in ten stakeholder focus groups at the end of the study. Participants included those who attended training in order to elicit their experiences and the impact of the two programs of education.

Data Analysis
Knowledge and skill performance ratings: Participant demographics, knowledge, and simulation skill performance were analysed using descriptive and inferential statistics. The relationship between paired knowledge data were explored using Paired Samples *t*-test and for ranked evaluation data (eg, competence, confidence). Wilcoxon Signed Ranks test. Independent samples *t*-test and ANOVA with post-hoc tests were used to test differences in performance scores. The Kruskal Wallis Test was used with dichotomous variables and Pearson correlation co-efficients were computed. The magnitude of the effect was interpreted using Cohen's recommendations (r effects: small ≥ .10, medium ≥ .30, large ≥ .50) (Cohen 1988).

Focus group analyses: Focus groups were conducted at the clinical sites from which the participants were recruited. The data was collected three months' post intervention (Sept 2017) aiming to elicit participant experiences and the impact of the two programs of education in managing deteriorating patients. Transcriptions were evaluated and analysed by two researchers using a coding format to cluster texts related to experience(s) in which language served to implicate aspects of practice. During the initial examination these experiences were reconstructed to enable the development of subjectivities that revealed tacit and embodied meanings derived from the various positions of the participants (Denzin & Lincoln, 2011).

Time Series Analyses
For *patient notes review:* a pre-post intervention analysis was conducted via time series analyses. The appropriateness of observations, incidence of early and late signs of deterioration (clinical review and MET calling criteria) were used as the primary measures of patient management.

Ethical Approval
Ethical approval was gained from the ethics committees at each of the hospital sites and approved by Federation University Australia (E15-018). Approval required participant consent and voluntary participation in the research components of the program.

RESULTS

1. Educational outcomes

Population: All qualified nursing staff (Registered and Enrolled Nurses) working in general medical wards were informed of the study and received an invitation to participate.

Response rate: 129 participants attended training (91% response rate) [LRH = 41/47; Sale = 24/32; SJOG Bendigo = 34/34; SJOG Berwick 27/28].

Participants completed the respective education program, following introductory onsite sessions, between February and June 2016. In the following section, results for each group are described separately.

Web-based program participants

Fifty-six (56) nurses from two hospital medical wards participated individually in the web-based simulation program FIRST^{2}ACTWeb. Some participant data was missed owing to an electronic collection error, however no substitutions were made. All but one were female, with age ranging from 21-62 years (mean 40, SD 11.9). Most spoke English at home 47/52 (90%). Qualifications included hospital training 18/50 (36%), a Bachelors degree in nursing 28/50 (56%), and a few held a post-graduate qualification 4/50 (8%). Clinical practice experience ranged from 1-43 years (mean 15.2±12.4) mostly in the acute medical field.

The median time taken to complete the program was 86 mins (range 23-382 mins), however one-fifth of participants completed in less than 45 mins. *[Note that it is possible start the program and return later to complete].* There was also a significant positive correlation between age and the length of time it took to complete (r= .36, p= 0.012). In other words older nurses spent longer on the program. Whether this was related to greater commitment or less familiarity with IT is unknown?

Knowledge development

Participants completed the MCQ before, and again after the program, with knowledge improving significantly (pre-test M= 7.56, SD 1.7; post-test 8.93, SD 1.7; (n=43) = t -5.990, p= <0.001) and with a medium effect size (r= .37). Eight of 12 knowledge items were significantly improved at post-test; in particular there was a marked improvement in an item that asked 'which are the six essential actions in the initial treatment of a deteriorating patient?. Bachelor-qualified nurses' knowledge improved the most (pre 7.59 SD 1.65; post 9.00 SD 1.29; z (n=22), -3.671, p=<0.001).

There was no correlation between participants' age and knowledge scores; however years of work experience was strongly negatively correlated with post-test knowledge scores (r= -.38, p= 0.03).

Skills Development

Participants completed the three screen based scenarios with notable increases in scores between scenarios 1-2, 2-3, and 1-3 (ANOVA, p= <0.05). Performances improved from 61% in the first scenario to 74% in the third scenario (Table 2).

Table 2: Scenario performance scores in web-based program

	1. Cardiac Scenario	2. Shock Scenario	3. Respiratory Scenario
Mean score	18.33, SD 4.30	19.15, SD 4.44	22.31, SD 3.24
Median score	18.00	20.00	22.00
Range	2-29	7-28	15-28
Maximum possible score	30	28	30
Mean score	**61.1%**	**69.6%**	**74.4%**

There was no significant correlation between participants' performance scores and length of experience, age, or qualification. However, performance scores were significantly correlated with post-test knowledge scores (*r*= .37, p = 0.014) with higher scores associated with higher knowledge levels.

Participant program evaluation

Seven evaluation questions were asked relating to the relevancy, applicability, level of challenge and the degree to which theory was integrated into practice. Participant evaluations were positive with an overall median rating of four from a possible five points with the program seen as appropriate, relevant, and stimulating. Suggestions for improvement included software limitations – not being able to carry out two tasks at the same time and no noticeable improvement in the patient's condition.

> *'In a clinical situation, I would normally be asking the patient their past history and how they were generally feeling while I was taking their vital signs. It felt like time was ticking away while I slowly did one action at a time.'*
>
> *'It was hard to know if the tasks I chose for management were effective until the end of each scenario.'*
>
> *'This program is very helpful in improving the knowledge on what actions to do on deteriorating patients. It improves critical thinking, management and teamwork within the multidisciplinary team.'*

Before and after the program participants self-rated their ability to recognise a deteriorating patient, to manage emergencies, and their level of confidence and competence. All ratings increased significantly (p=<0.001) with, for example, overall confidence increased (pre 2.96, SD 0.84; post 3.55, SD 0.68; z -3.938. p=<0.001) and with a medium effect size (*r* = .37).

Confidence ratings were also significantly correlated with years of working experience (*r*= .49, p= 0.001) and with skill performance scores (*r*= .50, p= <0.001). Competence ratings were also correlated with years working in healthcare (pre *r*= .42, p= 0.008) and with skills performance scores (*r*= .46, p= 0.020).

Face-to-face program participants

Seventy four (n= 74) nurses from medical wards in two hospitals participated in teams of three, in a clinical skills laboratory. The majority were female 69/74 (93%) with ages ranging from 22-65 (mean 37, SD 12.6). Most spoke English at home 59/68 (87%) and 23/74 (31%) were hospital trained, 47/74 (63%), held a Bachelors degree in nursing and 6/74 (8%) held a post-graduate qualification. Clinical practice experience ranged from 0-48 years (mean 12; SD 12.2 mostly in a medical or surgical field.

Knowledge development
Participants completed the MCQ before, and again after the program with knowledge improving significantly (pre-test M= 7.4, SD 1.6; post-test 8.1, SD 1.6; (n = 74) = t -3.991, p= <0.001) with a small effect size (r= .21). Six of 12 knowledge items were significantly improved at post-test, in particular a marked change in an item that asked 'which are the six essential actions in the initial treatment of a deteriorating patient?'.

Participants initial knowledge scores were significantly negatively correlated with age and years of working experience: younger nurses achieved higher knowledge scores (r= -.25; p = 0.034) as did less experienced nurses (r= -.25; p = 0.039). [n=70]. This pattern was repeated post-program showing a small negative correlation with knowledge and years of experience (r = -.26; p= 0.031).

Skills Development
There was no significant improvement in performance scores for the teams of nurses who completed scenario 1 (a mean score of 65.6%), scenario 2 (64.1%), or scenario 3 (65.8%).

Participant program evaluation
A set of seven evaluation questions were asked relating to the relevancy, applicability, level of challenge and the degree to which theory was integrated into practice. Participant evaluations were positive with an overall median rating of 5 from a possible five points (and a mean of 32.1/35) - with the program seen as appropriate, relevant, and stimulating. Participants' open-ended comments were positive:

> *'I learnt how to work through situations as a team and build confidence to care for patient before MET team arrives.'*
>
> *'Loved it! Oxygenation is vital and going back to nursing fundamentals is also important.'*
>
> *'I found the verbal feedback very helpful, thank you.'*

Before and after the program participants self-rated their ability to recognise a deteriorating patient, to manage emergencies, and their levels of confidence and competence. All ratings increased significantly (p=<0.001). For example overall confidence increasing by approximately 15% (pre 2.81, SD 0.79; post 3.55, SD 0.77; z -6.022. p=<0.001) and with a medium effect size (r = .43). There were no significant correlations with other variables such as age, length of experience, knowledge, or performance.

Overall educational outcomes
Knowledge: Objective measures of knowledge were reported in each cohort and showed that both programs improved participants' knowledge. When the increment was examined by

cohort, a non-significant difference was found between groups [WB: M= (39) 1.28, SD= 1.36. F2F: M= (74) 0.82, SD = 1.46. *t* = 1.619, df 111, p = .108 [CI= -0.10 - 1.02]. Thus both educational modalities were able to improve participants' knowledge equally.

Post-program evaluation: The post-program evaluation ratings showed that the F2F program received higher ratings on quality than the WB program, with six of seven question items rated significantly higher. For example F2F participants rated 'provision of effective feedback' (M= 4.62) significantly higher than those in the WB program (M = 4.15) (p = .005). This pattern continued with the F2F program being 'appropriate to my level of training' (p = .021) 'challenging but not threatening' (p= .001); 'enabling me to integrate theory' (p = .005); 'stimulated my interest' (p = .017) and 'encouraged me to think through a clinical problem. (p= .001). Overall, the summed evaluation ratings showed that the WB group ratings (M (n = 40) = 4.15 SD .89) were significantly lower than the F2F (M (n = 74) = 4.62 SD .66) *t*= -2.939, df 112, p= .005 [CI -.79 - -.15]. Thus the programs achieved significantly different feedback ratings, with F2F being rated significantly higher.

It was noted that there was no difference between training modalities in regard to participant post-program ratings of their 'ability to recognise or manage a deteriorating patient', nor their confidence or competence levels.

2. Clinical Impact

Patient records for each of the three months prior and the three months after the training intervention was audited.

Records audit:

In total, over 2,000 records were extracted, however a small number of these were found to be duplicated records (by ID and by same date of admission) and hence 1,951 records were pre-selected to remain in the dataset. The dataset was then filtered to remove admission records where a patient episode did not meet the inclusion criteria for the study (active medical treatment).

Excluded patients:

- Those under 18 years - n= 8.
- Patients with 'not for resuscitation orders' (DNR) - n=287 (14.7%).
- Patients with a DNR order with 'active treatment' limitations - n=23.
- Patients with an altered MET criteria - n= 64.
- Obstetric cases - n=5.

Thus, 19.8% cases were excluded from the final analyses. The remaining 1,564 patient records were analysed identifying that:

- 50.1% (n= 783) were admissions prior to the teaching intervention and 49.9% (n= 781) were after.
- Four research sites were represented, each site contributing between 21.5% and 29.0% of episodes (Table 3).

Table 3: Audit numbers by site

	Frequency	Percent
BERWICK	453	29.0
LATROBE REGIONAL HOSPITAL	374	23.9
BENDIGO	401	25.6
SALE	336	21.5
Total	1564	100.0

- The median age of patients was 69 (mean = 65.8, range 18-101); 45.8% were male and 54.2% were female.
- There was no difference in the pre- and post-sample by age; Pre M= 66, SD 18.65; post M= 65.7 SD 18.82 (p= 0.73) and no difference by sex.

Clinical Risk

Patient records indicated that 60.6% of all patients reached clinical review criteria during admission to the participating ward. Thirty percent (59%, n= 464) of patients reached clinical review criteria in the pre-intervention period and 62% (n = 482) in the post-test period. There was no significant difference between the pre and post samples by sample size (x^2 (1560) = 0.003, df 1, p= 0.96).

- The first day of admission carried the most risk for deterioration; 58.5% of documented first vital signs abnormalities occurred on the first day of admission, 27.6% on the second day, 7.0% on the third day and very few thereafter.

- Repeated charting of clinical risk observations in coloured risk sections of the observation chart occurred frequently. One-third of those reaching clinical review criteria (37%) were recorded at risk between one and three times during their admission; two-thirds were recorded between 4-8 times and the remainder up to a maximum of 116 times (any or all vital signs).

- There was some indication that charting was influenced by the day of the week or time of the day. For example, patients 'at clinical risk' were less likely to be charted as such on Sundays, overall rates were 5% lower (9.3% vs average 15% on weekdays).

Clinical Review

Of 464 patients pre- and 484 patients post training who reached clinical review criteria, 12% (n= 56) received a clinical review during admissions prior to the training intervention and 27% (n= 130) post-training. This indicates a significant positive improvement over all cases (x^2 (1564) = 0.003, df 1, p= <0.001) when considering those that reached clinical review criteria: (Mann-Whitney U Test: z (n= 926) = -6.001, p= <0.001). This was confirmed in time series analyses (TSA) which demonstrated a significant interventional effect (Figure 4).

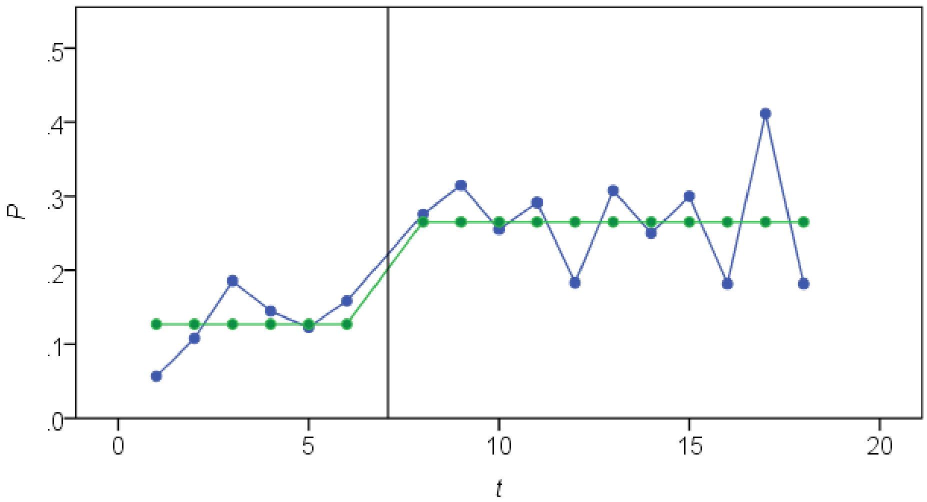


Figure 4: Clinical reviews: time series plot for values of P_1*across all four sites (blue) and fitted model (green) Vertical line indicates intervention commencement. *[*P_1 = the proportion of clinical review cases for which a review was initiated]*

- In summary the proportion of patients who received a clinical review increased significantly (more than doubled) after training.

Impact of nurses' training modality:
At pre-test in the WB cohort, the cases that reached clinical review criteria numbered 218; in the F2F group cases were n= 246. A Chi-square test of independence with continuity correction indicated there was no statistical difference between groups in the proportion that reached review criteria [x^2 (1, n= 781) = 1.97, p= 0.16, phi = -.05].

Further, whether at pre-test a clinical review was initiated for audited patients in the education groups was examined. A Chi-square test of independence with continuity correction indicated there was no statistical difference between groups in their decision to initiate a clinical review (x^2 (1, n= 465) = 0.12, p= 0.69, phi = -.02). Of an overall total n= 56 cases at pre-test who received a clinical review, there were n= 28/215 (13%) who were reviewed in WB and n= 28/246 (11.4%) in the F2F group.

- In summary at pre-test- There was no difference in the WB and the F2F training groups in the proportion of patients for whom they initiated a clinical review.

This between-group similarity remained at post-audit.

- At post-test - there was no significant difference between the WB and F2F group in the proportion of patients for whom clinical review was initiated (27% vs 28.8%). (x^2 (1, 465) = 0.11, p= 0.74, phi = 0.02). There were n= 58/215 for whom a clinical review was initiated in WB and n= 72/250 in the F2F group.

Summary: Although the proportion of patients for whom a clinical review was initiated increased significantly after training, there was no difference in the impact of the training technique on decisions to initiate a clinical review - between the WB and F2F groups i.e. both training modalities were equally as effective.

Medical Emergency Team calling criteria

Pre-intervention, 6.7% (n= 53) of all 783 patients reached MET calling criteria during admission. Post-intervention, 5.8% (n= 45) of all patients (n= 781) reached MET calling criteria. In summary there was no significant change between pre and post data sets regarding the frequency of patients who met the documented MET calling criteria (x^2 = 0.514, df 1, p= 0.473 (2-sided). However patients were less likely to deteriorate to MET calling criteria – 53 to 45.

Pre-intervention in the WB group there were significantly more cases where patients reached MET calling criteria (n= 38/384) compared with the F2F group (n= 15/399). A Chi-square test of independence (2x2) with continuity correction indicated there was a small statistical difference between groups (x^2 (1, 783) = 10.72, p= 0.001, phi = -0.12).

Pre-intervention for the WB group, a MET call was initiated for 12 patients (out of 38 patients - 31.5%)) and for the F2F group, for six patients (6 of 15 patients - 40%). A Chi-square test of independence (2x2) with continuity correction indicated there was no statistical difference in MET initiation between groups (x^2 (1, 53) = 0.07, p= 0.79, phi = 0.08).

MET activation post intervention

A MET team was activated in 33% (18/53) pre and 42% (19/45) post intervention with no statistically difference (p= .528).

Post-intervention- in the WB group there were significantly more cases where patients had reached MET calling criteria (n= 28/353; 7.9%) compared with the F2F group (n=17/428; 4%) (x^2 (1, 781) = 4.88 p= 0.03, phi = -0.08).

Post-intervention for the WB group, a MET call was initiated for 13 cases out of 28 pts (46%) and for the F2F, for six patients out of 17 patients (35%). This pattern appeared to follow the pre-test cases where there were higher levels of MET calls. However, there was no statistical difference in MET initiation between groups after training (x^2 (1, 45) = 0.18, p= 0.67, phi = -0.11). These data may be limited by the small numbers in each group which limits statistical computations that may have identified practice differences.

Further, after training- for the WB group there was no significant change in the proportion of patients for whom a MET call was initiated (p= 0.223). Similarly, for the F2F group after training there was no significant change in the proportion of patients for whom a MET call was initiated (p= 0.823).

Escalating Care

- Pre-intervention, 8 patients (1%) were admitted to ICU and post-intervention, 10 were admitted to ICU (1.3%).

- Median length of stay in ICU was four days (Mean 4.2, SD 2.33, range 1-44).

Nursing Documentation

Multiple variables were collected to explore documentation practices. Overall results are presented in Appendix 1. Where patients met clinical review criteria vital signs were almost always recorded. Clinical reviews criteria breaches were most likely to be met for altered SpO2 levels (51.5% of all patients), blood pressure (53.3% of all patients) and temperature (40% of all patients).

TSA identified a decline in inappropriate frequencies of vital sign recording across both the pre and post-intervention periods indicating that the interventions continued the quality improvement process (Figure 5).

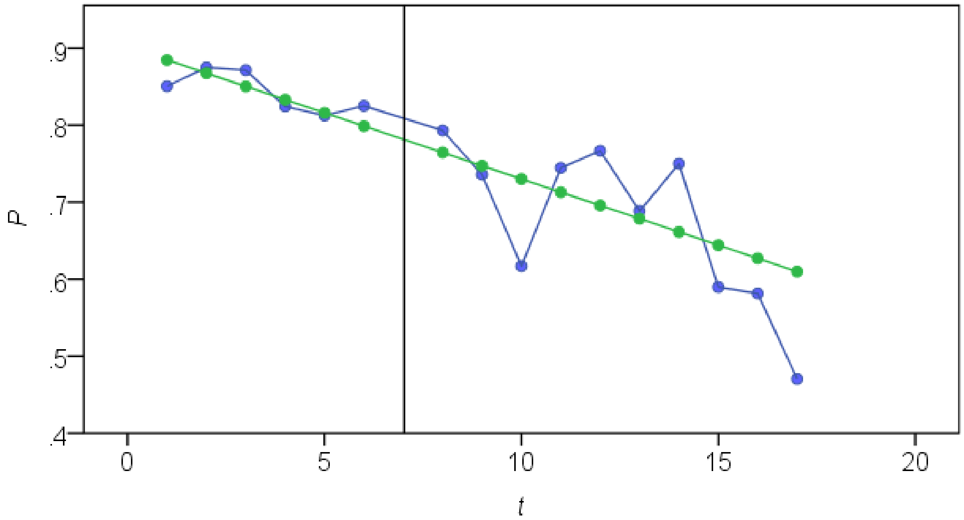


Figure 5: Vital sign recording: time series plot for values of P_3*across all four sites (blue) and fitted model (green) Vertical line indicates intervention commencement. *[*P_3 = the proportion of inappropriate frequencies of vital sign recording]*

TSA identified that inappropriate oxygen therapy (e.g. nasal prongs instead of a Hudson mask) declined pre-intervention and in the first two months post intervention, but was not maintained beyond this point (Figure 6).

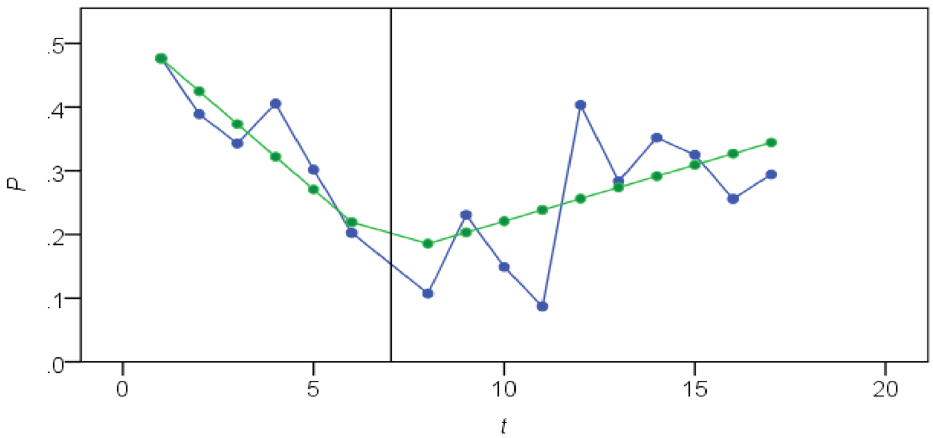


Figure 6: Oxygen therapy: time series plot for values of P_4*across all four sites (blue) and fitted model (green) Vertical line indicates intervention commencement. *[*P_4= the proportion of inappropriate oxygen therapy]*

TSA identified a significant interventional effect on the applicable recording of pain scores with a significant reduction in cases where pain score was not recorded (Figure 7).

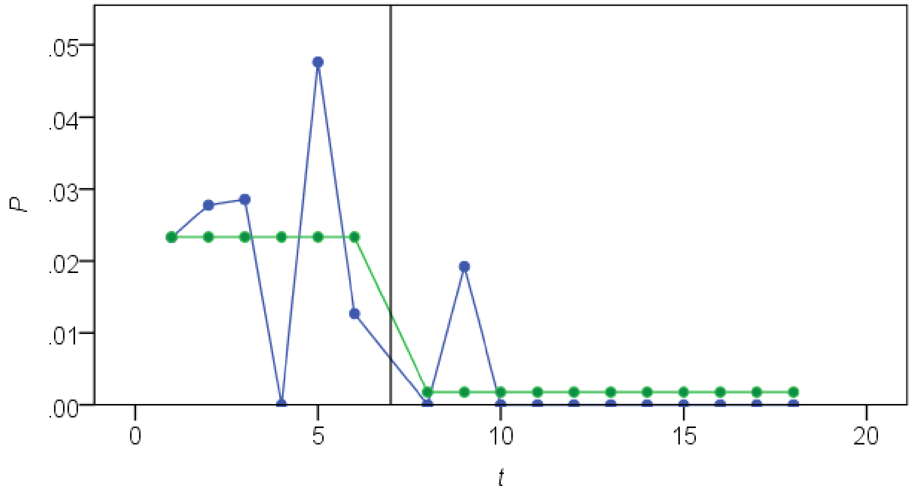


Figure 7: Pain scoring: time series plot for values of P_5*across all four sites (blue) and fitted model (green). Vertical line indicates intervention commencement. *[*P_5= the proportion of pain scores that were not recorded]*

In addition, post intervention, there was an increased recording of conscious state (AVPU) (by 8.9%). Further, more 'at-risk' patients were reviewed within 60 mins and more interventions were performed - post intervention (p= <0.05).

Where patients met MET calling criteria there was an improvement pre-post intervention in the applicable delivery of oxygen. MET criteria were most likely to be met for altered SpO2 levels, heart rate and blood pressure.

3. Economic Analyses

A cost analysis was conducted to evaluate the difference in costs between WB training and F2F training in nurses' recognition and management of deteriorating patients (Cooper et al, 2016). The main cost elements evaluated were the training costs associated with each training intervention. However, to support the validity of the records audit we also modelled the similarity/difference in hospitalization costs between those patients cared for by the respective nursing groups based on our patient records audit data. Hospitalization costs were calculated using the AR-DRG costing system and resource intensity weights (WEIS 2016-17)[3]. Hospitalization cost calculations are described in Appendix 2. In addition we collected the LOS data for 2015 and 2016 for all patient in the participating wards.

Training Intervention Costs

The overall training costs over the 6 month study were greater in the WB group ($56,514; mean cost $1027/participant) compared to the F2F group ($36,276; mean cost $490.22/participant) by a factor of 1.5:1.0. (*This includes a one-time software development cost of $32,969 for the WB program)* (see Appendix 2).

As there were differences in the number of nurses trained in the F2F group (n= 74), and the WB group (n=55) over the study period, a sensitivity analysis was performed whereby total and mean training costs for each group are modeled with 1) an equivalent number of trainees and 2) if additional nurses were trained (Tables 4 and 5).

- Overall, as the number of trainees increase, the difference in total training costs between the two groups decreases. At the point where n= 228 nurses are trained, the total costs between the WB and F2F groups are equal. Above this number of trainees, the total and mean cost of the WB intervention becomes less than F2F.
- The cost of training 300 participants was $240.52 per WB trainee and $276.75 per F2F trainee.

Table 4: Sensitivity analysis- Projected training cost for Web-based training group

Cost Type	Criteria	Trainee=100 (6M)	Trainee=200 (12M)	Trainee=300 (18M)
Project Management	One time	19,515	19,515	19,515
Average Trainee cost	Continuous	56.15*100=5,615	56.15*200=11,230	56.15*300=16,845
Trainer	Continuous	942*1=942	942*2=1,884	942*3=2,826
Filming	One time	8,169	8,169	8,169
Software	One time	24,800	24,800	24,800
Total costs ($)		$59,041	$65,598	$72,155
Mean costs of training		$590.41	$327.99	$240.52

Assumptions:

1. Average trainee cost is based on 1.5 hours of training time for each trainee.
2. The trainers are assumed to train 100 participants in every 6 months; therefore adding an extra training session for each 100 participants.
3. All costs are reported in AUS $.

Table 5: Sensitivity analysis - Projected training cost for Face2Face training group

Cost Type	Criteria	Trainee=100 (6M)	Trainee=200 (12M)	Trainee=300 (18M)
Project Management	One time	15,910	15,910	15,910
Average Trainee cost	Continuous	77.16*100=7716	77.16*200=15432	77.16* 300=23148
Trainer	Continuous	14,656*1=14,656	14656*2=29,312	14656*3=43,968
Total costs ($)		$38,282	$60,654	$83,026
Mean costs of training		$382.82	$303.27	$276.75

Assumption:

1. Average trainee cost is based on 2 hours of training time for each trainee.
2. The trainers are assumed to train 100 participants in every 6 months; therefore adding an extra training session for each 100 participants.
3. All costs are reported in AUS $.

Web-based Intervention

Overall mean LOS and hospitalization costs were higher in the post training period (6.77 days and $6,928.40 respectively) compared with the pre training period (5.85 days and $6,040 respectively) (Appendix 2). The mean LOS for all patients in the two participating wards during approximately the same periods was 6.85 pre training and 6.75 post. In the previous year (2015) for the same periods, this was 6.50 pre training and 5.40 post. Overall patient intensity, based on the resource intensity weight assigned to the respective patient DRG, did not appear to contribute to the higher post intervention hospitalization costs. There was no difference in the number of individuals who went to ICU or their length of stay in ICU pre or post intervention in the web-based training group.

Face2Face Intervention

Overall mean LOS and hospitalization costs were slightly higher in the post training period (5.16 days and $5,302.10 respectively) compared with the pre training period (4.77 days and $5,280.60 respectively) (Appendix 2). The mean LOS for all patients in the two participating wards during approximately the same periods was 4.35 pre training and 4.90 post. In the previous year (2015) for the same periods this was 4.65 pre training and 4.90 post. No difference in patient intensity in the pre and post intervention period was observed. There was no difference in the number of individuals who went to ICU pre or post intervention however the length of stay in ICU post intervention was slightly less than in the pre intervention period in the F2F training group (3.2 versus 5.75 days).

Hospitalisation outcome

Comparing the difference in mean hospitalization costs in the WB group to the mean hospitalization costs in the F2F group, the hospital costs for the group who were trained using the web-based intervention is on average $858.30 higher. However, the variation around this mean difference is quite high and is therefore not considered statistically significant (95% CI using bootstrap sample -$973 to $2648). Note also that the LOS for all ward patients in these two WB trained hospitals was notably higher than for the F2F group in both 2015 and 2016. [e.g. 2016 - WB pre 6.85, post 6.75 – F2F pre 4.35, post 4.90].

We conclude that the hospitalization costs are hypothetical as they are unadjusted for several variables which would influence costs (difference in patient acuity/clinical condition/seasonal variations between hospital sites) and therefore the training groups. These were outside the scope of this analyses. Based on the resource intensity weights (these are multiplied by a base hospital rate to get the overall hospital cost for a specific DRG patient) between WB and F2F groups, the difference may be due to an overall less resource intense ward population in the F2F group compared to the WB group.

Overall, from the bootstrap analyses there was no significant difference in LOS between the WB and F2F intervention groups (0.51 days (95%CI (-0.62, 1.56). However the raw data (Appendix 1) indicates a LOS of 6.31 for the WB ward sample and 4.94 for the F2F ward samples – notably higher in the WB hospitals - which corresponds with the whole ward samples for 2016 and 2015 in the same periods. Although there was a difference in mean costs in the pre and post WB group with higher post intervention hospitalization costs, there was no overall difference in mean cost differences between WB and F2F training during the study period ($858.30 (95% CI (-973, 2648). In conclusion a range of factors influence costs and it is unlikely that the different training interventions had different influences on hospitalization costs, especially as the LOS differed between the WB and F2F hospitals.

4: Focus group findings

Participants who had completed either of the education interventions were invited to take part in the focus group discussions. Ten focus groups were held [SGOG, Berwick -4 groups 32 participants; LRH, Traralgon - two groups, 13 participants; CGHS, Sale, - two groups, 6 participants, SJOG, Bendigo, - two groups 15 participants] three months post intervention (Sept 2017) aiming to elicit participant experiences and the impact of the two programs of education. In total 66 participants attended designated as graduate nurses (n=7) (1st year post registration) enrolled nurses (n=18) registered nurses (n=31), associate nurse unit managers (ANUM) (N=6) and general managers/educators/learning program coordinators (n=4). Focus groups lasted approximately 1 hour and were audio recorded and transcribed. During the focus groups semi-structured questions were used to invite commentary on experiences of:

1. the form of program completed (i.e. face-to-face or web based).
2. the fidelity of the simulations.
3. the clinical applicability of the program.
4. educational outcomes.
5. how the program could be improved.
6. reflections on prior and future practice.

Transcripts from the focus groups were analysed using a coding format to cluster texts related to experience(s), meanings and aspects of practice that were then grouped into sub and common themes (Denzin and Lincoln, 2011). Lessons learned from these programs are outlined in the context of participants' current experience and the potential for influencing the future of education in the field. In the following summary specific themes are reflected with participants' thoughts. Core themes were '**Program impact'; 'Managing deterioration'** and **'Clinical judgment'.**

1. **Program impact**

Participants from both programs identified key outcomes of the experience including a) the curricula, b) debriefing, c) limits of simulation, d) clinical impact, e) anxiety, and f) problems solving.

1a The curricula

Comments relating to the web-based program varied relating to access, collaboration and technology:

> *"it was good. Online learning, I like that business. Because you can do it at your own pace."(wb).*
>
> *It shows, it prioritises what are things has to do. Also the other thing when we do that, we when we get a good score at the end, it actually increases the confidence (wb).*
>
> *"…with the online version, is there's no room there for discussion and to clarify and those sorts of things that you can in a face-to-face presentation that's…" (wb).*

"less collaborative than the actual - than the event actually is in a hospital. It's much more collaborative"(wb).

"a lot of us are a little bit older, less techno savvy, and perhaps learn face-to-face better than –"(wb).

Whilst participants in the face-to-face program noted the generic nature of the training and the public/private hospital divide:

'It's a few fundamentals that I think is quite different between private hospitals and public hospitals; so how we work wasn't really differentiated. Yes, the scenarios would all be the same, but the way we work within the hospital was very different."(f2f)

"I think it's more perhaps suited for a public hospital than a private. Like I work public and private and it was very - probably more like what would happen in a public hospital, having the doctor right there and on standby. Whereas here you don't."(f2f)

However both programs were thought to be good refreshers and a great way to develop and update:

"I found it's just been a bit more reinforcing. I've been nursing for a hell of a long time so it was more reinforcing".(f2f)

"I thought it might be good if we do 12 month mandatory education with fire and safety and forms and all that sort of stuff. But I thought it was a good refresher of like a medical refresher. What you should do when you should see that deterioration, just a refresher that sometimes you might not notice something that you should."(wb)

1b Debriefing

Many comments were made about the feedback and debriefing process in both programs:

"It also builds your confidence to know that you actually have got that knowledge in there and that you are doing the right thing when you get your feedback."(wb)

"I thought it was a positive experience. It was a powerful learning experience. So to be able to self-reflect afterwards and yeah, talk about what we could have done and what we didn't do and what we did do."(f2f)

"It's different when you're in the situation, you just go and do things. Whereas being outside and then looking at the film was another experience that you could pick up on things that in the moment - that you weren't necessarily properly concentrating on or thinking of. So it just gave you another perspective being outside, looking in and not part of the situation."(f2f)

However debrief limitations relating to resourcing, time constraints and IT constraints were also raised:

"There was no 'What I should have done better?', or 'What I could have done better?' because it was kind of like, 'This is what you've done, and 'This is what you think you

should have done.' But was it actually right? I didn't know. So yeah, I would have liked a little bit more follow-up; perhaps next time – 'Do this, and this' ".(f2f)

"I remember at the time feeling frustrated that there wasn't more feedback". (f2f)

"Going back to my earlier comments about feedback, I think that would be the most valuable aspect of the whole training. As I said before, I felt that the scenarios were excellent and made us think quickly, which is really important. It's ultimately having a little bit more feedback surrounding our actions and areas to improve would have been ideal because, I guess, that validates the process and highlights any deficits et cetera".-(f2f)

1c- Limits of Simulation

Participants raised concerns around the limits of the simulations – in the face-to-face program this related to an actors inability to mirror vital signs and the constraints of allocating 'doctor' in each scenario who was unable to prompt or initiate action:

"when they're an actor - you have to actually ask are they sweating, are they clammy That makes it obviously not real to me, anyway"(f2f)

'I wouldn't expect the staff to be prompting the doctor, whereas that's what it felt like" (f2f)

"Because you knew we were hands on and we were all straight in there and we just found it weird that this doctor said ' I don't know', 'I don't know', you know what the answer was but can we tell you what it is. Because there's a camera there and we're all looking at each other, and yeah, it's a bit ….. (f2f)

Web based participants raised concerns about their inability to multi-task and that patients' condition continued to deteriorate despite treatment:

"It was frustrating because if you're in a clinical situation you can do two things at once. While you're, say, taking an ECG you'd be asking the patient about their family history or anything like that. You couldn't do that. You had to specifically do one thing and then you're locked out for that length of time, which seemed, when the clock's ticking, it seemed like a long time. So I found it very frustrating. "(wb)

"The patient didn't change at all, no matter what I did. So I'm thinking I must have to do more."(wb)

1d – Clinical Impact

Participants also commented on how the programs had changed their clinical practice – emphasising the need for improved documentation and 'triggering' calls for help:

"I've become more aware of my documentation, which is hard in constraints of time."-(f2f)

"Just document, document everything. Even if someone else says, no, don't worry about it, just make sure you've done it." (f2f)

"I'm definitely thinking more about the MET call situation now. If there isn't something done already that I can do, do it rather than just standing there and waiting…"(f2f)

"I think it also helped identify the gaps on the observations sheets and what we were and weren't doing correctly. I think I've even seen now back on the ward that people are utilising them … the obs charts now…. and documenting it properly." (f2f)

"That program shows me how important the frequency of doing the observations and stuff, doing a MET call thing and also it's really helpful when we get that second or third hand on time. When we call that MET call on time." (wb)

Some also commented on how the face-to-face program had enhanced problem solving skills:

"I think it helped the problem solving process, yeah. I think it facilitated that and sort of helped guide you through what you'd sort of do mentally anyway" (f2f).

1e – Anxiety

Both programs also generated a degree of anxiety which may have enhanced performance or hindered it?

"People who weren't computer literate were more anxious."(wb)

"It made you feel a little bit anxious, because you were doing everything on your own. You didn't have that support you felt." (wb)

'The scenarios were fine. I just didn't enjoy the experience. I don't like being filmed. I don't have photos taken, so the actual thought of it being filmed was quite horrific".- (f2f)

2- Managing deterioration

Concerns about the management of deteriorating patients were frequently voiced and related to communication issues:

"A good handover is key. So it's making sure that that continuity in communicating is always open and that information is passed on."(f2f)

"Clear expectations. C*lear guidelines and expectations from the medical team would absolutely go a long way" I think if there's clear direction it eliminates all of that uncertainty. (f2f)*

Documentation of altered MET calling criteria was also a concern:

"Someone's deteriorating and there's been a decision that they won't take any further action but it hasn't been clearly documented. "(wb)

"Altered MET, that hasn't been clarified." (wb)

"There's quite a few altered conscious states that do or don't get called. In different situations. People that are in palliative care situations for instance that have still MET calls that that hasn't been - they've been called but hasn't been altered yet. They're in that phase that we wouldn't call MET call."(wb)

Help seeking behaviours varied and were often contrary to vitals sign charts criteria for triggering assistance:

"You might get the CCU liaison to come around and then you've got your buddy helper and you might manage it before it gets to the MET call." (f2f)

"We were doing everything that needed to be done and we transferred the patient out. We didn't need to do all the bells and whistles". (f2f)

"Being told not to (call for help). *If you run it past your supervisor or whoever's in charge and they say no. Nothing you can do about it." (wb)*

3- Clinical Judgement

There was striking results in relation to the use of the prescribed vital signs charts with strong consensus that compliance reduced individuals' critical thinking flexibility. Freedom to make clinical judgments was a concern:

"Sometimes - we're not encouraged to have clinical - or have clinical judgement. It says that we've got this, our clinical judgement is not as appreciated. So therefore, - you don't develop it, As soon as it hits those colours, where we focus on calling a MET, but we don't focus on our deteriorating patient, if you know what - does that make sense? (f2f)

Confidence – managing the whole:

"We're not also thinking of why is it going down, we're thinking oh, it's hit the yellow. I just need to call. We're not thinking well, it's going down. They're a bit - I think they might be a bit dry. We need to think of what interventions that we need to do at that stage."(wb)

"It's about the management which - that reiterating that we're managing and what we need to do and what's right rather than we've called the MET, someone else's responsibility." (wb)

"I reckon the nurse's clinical call often is the right call. Rather than just relying on the raw numbers". (wb)

The importance of intuition but the fear of retribution:

"If you're concerned about a patient. I mean, regardless of - I would say, for me, it would be gut instinct. If I'm concerned and the obs look okay - sometimes the observations aren't really going to be your prompt to call." (f2f)

"Sometimes we might be discouraged from making that right decision when we should because we've been bitten a couple of times where we've rang them" (wb)

Discussion

This study explored the contribution of simulation-based education programs for ward based nurses aiming to improve recognition and response to patient deterioration. As per Standard 9 of the Australian National Safety and Quality Standards, health services are required to have systems and processes in place regarding: *Recognising and Responding to Clinical Deterioration in Acute Health Care* (ACSQH, 2012). As nurses in general wards are often the primary responder (O'Leary et al. 2015) knowledge, skills and situation awareness are required to respond effectively to patients when the patient's clinical condition deteriorates (Kinsman et al, 2012).

Clinical practice improved at all four sites after the training interventions. Interventions to review a patient when they reached clinical review criteria more than doubled (12.2% to 28%) which was confirmed in time series analyses (TSA) as a significant interventional effect. This will have a major impact on patient safety through the timely and appropriate review and care of patients with early signs of deterioration. The treatment of the remaining 72% who had no recorded review is unknown, and is cause for concern. The 'failure to rescue' literature clearly describes the deficit in escalating clinical care in deteriorating patients (Kitto et al, 2015). Some reasons why nurses would not actively treat a patient who was at risk were given in focus group discussion- for example undocumented altered alert criteria.

Further, detailed time series analyses identified that inappropriate oxygen therapy (e.g. nasal prongs instead of a Hudson mask declined pre-intervention and in the first two months post intervention, but was not maintained beyond this point. This may have related to differing cultural norms within the wards and confusion over oxygen therapy guidelines? TSA also identified an applicable increase in the frequency of vital signs recording across both the pre and post-intervention periods indicating that the interventions continued the quality improvement process. Participants were made aware that training was due prior to commencement which may have heightened their awareness of best practice in advance of the interventions. Finally TSA identified a significant interventional effect on the recording of pain scores, more 'at-risk' patients were reviewed within 60 mins and more interventions were performed - post intervention.

Using objective assessments to measure the WB and F2F program educational outcomes knowledge significantly improved in both participant groups ($p<0.05$). This finding aligns with our prior studies where the knowledge of registered nurses was statistically improved after face-to-face training (Kinsman et al, 2012) as was the knowledge of students (Buykx et al 2012) and the knowledge of students in WB training Sparkes et al, 2016).

In line with a preference for use of valid objective measures to enhance the quality of program evaluations (Cant & Cooper 2016) skill performances in the three scenarios were rated via either face-to-face observation, or in the case of WB by an online computer-based scoring system. Over the three scenarios the WB program showed that participants' skill performance improved, however there was no change to the F2F performance ratings. This underlines the differences between simulation modalities in the way the programs were delivered and aided learning. In this study participants in the F2F program only led a scenario once and did not have the repeat individual practice enabled in the WB program. It was also apparent from focus group findings that nurses were apprehensive about their clinical performance as they were mostly unfamiliar with simulation-based training involving active role play. The theme of ‘anxiety’ was identified in the focus groups and nurses’ confidence was initially low but increased significantly by the end of the program. This improvement in confidence may have been due to the detailed debriefing provided in the F2F program and why this modality was more highly evaluated than the WB program. We suggest therefore that repeated role-plays, using simulation scenarios, are an ideal way to improve nurses’ confidence which may then be transferable to clinical practice.

Demographic factors also appear to play a part in outcomes. For example older nurses spent longer on the WB program, which may be related to greater commitment or less familiarity with information technology. Further, younger less experienced participants in the WB and F2F program gained higher knowledge scores.

An economic analyses identified that the WB version costs more to implement initially (including the cost of program development) but reduces significantly over time in comparison to F2F training. For example the cost of training ‘the first’ 100 WB participants is $590 per person. By the time 300 participants have been trained the costs have dropped to $241 per person (a 59% reduction). This is compared to the cost of training ‘the first’ 100 F2F participants at $383 per person which by 300 participants drops much less to $277 (a 28% reduction. Further, in relation to hospital costs there was no significant difference in the mean patient hospitalisation costs related to the two interventions.

In summary therefore and in response to our research questions – both models of training did improve the detection and management of deterioration patients. Although the WB program was more costly to implement over this study period, the overall and mean costs of the WB program becomes less than the F2F as the number of trainees increase.

Limitations and Issues

The findings are limited in that the interventions were completed in a relatively small sample of nursing staff in four hospitals in Victoria, Australia. This does therefore limit the generalisability of the findings internationally. Further, there were some data collection issues with some WB participant data missing due to an electronic data collection error. Patient survey data was also collected electronically through Qualtrics and we found that a number of patient cases had been repeated when there was a web connection fault. These were easily identified and deleted but such issues should be considered when using electronic means of data collection – which may be faster but is still prone to error.

In Summary

- Both WB and F2F programs were positively evaluated.

- Both WB and F2F programs significantly improve knowledge, confidence and competence.

- WB enhances skill development whilst F2F training is likely to improve skills with repetitive practice.

- Both programs had a significant clinical impact, increasing clinical reviews and improving nursing interventions.

- Upfront costs are higher in the WB version compared with F2F however over time implementation of the web based version will be considerably less costly.

- Interviews indicate positive educational and clinical impact but room for improved debriefing, limits to the fidelity of simulation, and a need to consider clinical communication, vital sign chart forms and practice.

The future

- To improve nurses' skills in detecting and managing patient deterioration, the use of both programs consecutively is recommended i.e. 'Blended learning'.

- Bearing in mind the continued poor compliance to vital signs recording/interventions there is a need to review charts and response criteria.

- Expansion of these simulation programs to inter-disciplinary teams would add to our understanding of how simulation can contribute to better patient outcomes in inter-disciplinary environments.

References

Australian Commission on Safety and Quality in Healthcare (2012). National Safety and Quality Healthcare Standards Canberra.

Buykx P. **Cooper S. Kinsman L.** Endacott R. Scholes J. McConnell-Henry T. **Cant R.** (2012) Patient deterioration simulation experiences: impact on teaching and learning. *Collegian*. 19 (3). 125-129

Buykx, P., **Kinsman, L, Cooper, S,** McConnell-Henry, T, **Cant, R**, Endacott, R, Scholes, J., (2011). FIRST2 ACT: educating nurses to identify patient deterioration – a theory-based model for best practice simulation education. Nurse Education Today 31(7): 687-693 DOI: 10.1016/j.nedt.2011.03.006

Cant R. P., Cooper S. J. (2016) Use of simulation-based learning in undergraduate nursing education: An umbrella systematic review. *Nurse Education Today*, 49: 63-71. Doi: http://dx.doi.org/10.1016/j.nedt.2016.11.015

Cioffi, J. (2000).Nurses' experiences of making decisions to call emergency assistance to their patients. Journal of Advanced Nursing 32: 108-114.

Cioffi, J. (2001). Clinical Simulations: Development And Validation. Nurse Education Today, 21, 477-486. 21: 477-486.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates

Connell, C. J., Endacott, R., Jackman, J. A., Kiprillis, N. R., Sparkes, L. M., **Cooper, S. J.**, (2016). The effectiveness of education in the recognition and management of deteriorating patients: A systematic review. Nurse Education Today 44: 133-145 doi:10.1016/j.nedt.2016.06.001.

Cooper, S., Kinsman, L, Chung, C, Cant, R, Boyle, J, Bull, L, Cameron, A, Connell, C, Kim, J-A, McInnes, D, McKay, A, Nankervis, K, Penz, E, Rotter T. (2016). The impact of web-based and face-to-face simulation on patient deterioration and patient safety: protocol for a multi-site multi-method design. BMC Health Services Research 16(1): 475. doi:10.1186/s12913-016-1683-0.

Cooper, S., Kinsman, L., Buykx, P., Mcconnell-Henry, T., Endacott, R. & Scholes, J. (2010). Managing The Deteriorating Patient In A Simulated Environment: Nursing Student's Knowledge, Skill And Situational Awareness. 19, 2309-2318." Journal Of Clinical Nursing 19: 2309-2318.

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. Thousand Oaks: Sage.

Endacott, R., Scholes, J., Freeman, M. & **Cooper, S.**, (2003). The Reality Of Clinical Learning In Critical Care Settings: A Practitioner: Student Gap?" Journal Of Clinical Nursing 12: 778-785.

Endacott R. **Cooper S.** Scholes J. **Kinsman L**. McConnell-Henry T. (2010) Final-year nursing students' ability to assess, detect and act on clinical cues of deterioration in a simulated environment. *Journal of Advanced Nursing*. 66(12), 2722–2731.

Endacott, R., Westley, M. , (2006). Managing patients at risk of deterioration in rural hospitals: A qualitative study. Australian Journal of Rural Health 14: 275-79.

Fritz, P., Flanagan, B., (2008). Review of mannequin-based high fidelity simulation in emergency medicine. Emerg Med Aust 20(1-9).

Glick, H., Doshi, JA, Sonnad, SS, Polsky, D., (2007). Economic evaluation in clinical trials. Oxford, Oxford: University Press.

Harrison, G., Jacques, TC., Kilborn, G, McLaws, M., (2005). The prevalence of recordings of the signs of critical conditions and emergency response in hospital wards – the SOCCER study. Resuscitation 65: 149-167.

Hogan, M. P., Pace, D. E., Hapgood, J. & Boone, D. C., (2006). "Use Of Human Patient Simulation And The Situational Awareness Global Assessment Technique In Practical Trauma Skills Assessment. ." Journal Of Trauma- Injury Infection And Critical Care 61: 1047-1052.

Kinsman, L., Buykx, P, Champion, R, **Cant, R, Cooper, S,** Endacott, R, McConnell-Henry, T, Porter, J, Missen, K, Scholes, J, (2012). "The First2Act simulation program improves nursing practice in a rural Australian hospital." Australian Journal of Rural Health 20(5): 270-274.

Kitto, S., Marshall, S. D., McMillan, S. E., Shearer, B., Buist, M., Grant, R., . . . Wilson, S., (2015). Rapid response systems and collective (in)competence: An exploratory analysis of intraprofessional and interprofessional activation factors. Journal of interprofessional care 29 (4), 340. doi:10.3109/13561820.2014.984021.

Liaw, S. Y., Wong, L. F., Lim, E. Y. P., Ang, S. B. L., Mujumdar, S., Ho, J. T. Y., . . . Ang, E. N. K., (2016). Effectiveness of a Web-Based Simulation in Improving Nurses' Workplace Practice With Deteriorating Ward Patients: A Pre- and Post intervention Study. Journal of Medical Internet Research 18(2): e37. doi:10.2196/jmir.5294.

McConnell-Henry T; **Cooper S,** Endacott R. Porter J. Missen K. Sparkes L. (2015) Designing a medical records review tool: An instructional guide. Contemporary Nurse. 50. 1. 72-79 DOI: 10.1080/10376178.2015.1010260

Motola, I. and L. Devine (2013). Simulation in healthcare education: A best evidence practical guide. AMEE Guide No. 82. Medical Teacher 35: e1511–e1530.

Needleman J (2013) Increasing acuity, increasing technology, and the changing demands on nurses. Nursing Economics. 31(4):200-202.

O'Leary, J., Nash, R., & Lewis, P., (2016). Standard instruction versus simulation: Educating registered nurses in the early recognition of patient deterioration in paediatric critical care. Nurse Education Today 36: 287-292. doi:http://dx.doi.org/10.1016/j.nedt.2015.07.021.

Ramsey, S., Willke, R, Briggs, A, et al. (2005). Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health 8: 521-533.

Scholes, J. Endacott R, Bulle B, **Cooper S**, Miles M, Gilmour C, Buykx P, **Kinsman L**, Boland R, Jones J. Zaidi F. (2012). Clinical decision-making: midwifery students' recognition of, and response to, post partum haemorrhage in the simulation environment. BMC Pregnancy and Childbirth 12: 19.

Sparkes L. Chan M. **Cooper S**. Pang M. Tiwari A. (2016) Enhancing the management of deteriorating patients with Australian on line e-simulation software: Acceptability, transferability and impact in Hong Kong? Nursing and Health Sciences. 18; 393-399 .

Waldie, J., Day, T., Tee, S., (2016). Patient safety in acute care: are we going around in circles? British Journal of Nursing 25(13): 747. doi:10.12968/bjon.2016.25.13.747.

Ziv, A., S. Ben-David, Ziv, M. (2005). "Simulation based medical education: an opportunity to learn from errors." Medical Teacher 27(3): 193-9.

Appendix 1: Nursing Documentation

Recording of vital signs

		PRE-INTERVENTION [N=464] n/%	POST-INTERVENTION [N=484] n/%	
CLINICAL REVIEW – Vital signs recorded				
Date/time on obs chart	Yes	454 (97.8)(	451 (95.8)	
Context SpO2 : Recorded (e.g. 99% on RA)	Yes No	399 (87.7) 56 (12.3)	460 (95.3) 13 (2.7)	
Supplemental O2 delivered applicably:	Yes No	296 (64.5) 163 (35.5)	346 (73.0) 128 (27.0)	
Supplemental O2 : Device used	NP Hudson Mask Non-rebreather	83 11 0	77 15 3	
Was SpO2 recorded?	Yes	465 (99.8)	479 (100.0)	
Was the Temperature recorded?	Yes	463 (99.8)	477 (99.6)	
Was Pain Score recorded?	Yes	452 (97.8)	478 (99.8)	
Was HR recorded?	Yes	460 (96.6)	479 (100.0)	
Was RR recorded?	Yes	459 (99.6)	477 (99.8)	
Was BP recorded?	Yes	460 (99.8)	476 (99.8)	
Was AVPU recorded?	Yes	457 (91.1)	475 (100.0)	
Was BGL recorded?	Yes	137 (29.8)	108 (22.9)	
Clinical Review Criteria reached?		N= 464 patients	N= 484 patients	

Vital signs for which clinical review was reached				
SpO2 criteria?	Yes	248 (53.4)	237 (49.6)	
Temperature criteria?	Yes	157 (33.8)	219 (46.3)	
Pain Score criteria?	Yes	2	14 (3.0)	
HR criteria?	Yes	83 (17.9)	94 (20.1)	
RR criteria?	Yes	19 (4.1)	37 (8.0)	
BP criteria?	Yes	224 (48.2)	275 (58.4)	
AVPU criteria?	Yes	17 (3.6)	24 (5.2)	
BGL criteria?	Yes	13 (2.8)	10 (3.0)	
Did the frequency of observations increase when clinical review criteria was reached?	Yes	71 (15.3)	145 (30.0)	
Was the patient reviewed within 60 minutes?	Yes	52 (11.2)	129 (26.7)	
Were any intervention/s administered and recorded upon meeting clinical review criteria?	Yes	98 (21.1)	164 (33.9)	

MET - Vital signs recorded				
		N=53 patients	N=45 patients	
Date/time on obs chart	Yes	49 (92.5)	43 (95.6)	
Context of recorded SpO2 (e.g. 99% on RA)	Yes	44 (83.1)	41 (91.1)	

Supplemental O2 : delivered applicably?	Yes	34 (66.0)	37 (84.1)	
Supplemental O2 : Device used-Supplemental O2 delivered?	NP Hudson mask non-re-breather	13 4 0	9 4 2	
Was SpO2 recorded?	Yes	50 (94.3)	41 (95.3)	
Was the Temperature recorded?	Yes	47 (88.6)	41 (95.3)	
Was Pain Score recorded?	Yes	47 (88.6)	40 (93.0)	
Was HR recorded?	Yes	48 (90.5)	43 (100.0)	
Was RR recorded?	Yes	46 (86.8)	42 (97.7)	
Was BP recorded?	Yes	52 (98.1)	43 (100.0)	
Was AVPU recorded?	Yes	41 (77.3)	32 (74.4)	
Vital signs for which MET calling criteria was reached				
SpO2 criteria?	Yes	20	[N=41] 15 (30.7)	
Temperature criteria?	Yes	3	2 (4.9)	
Pain Score criteria?	Yes		2 (5.0)	
HR criteria	Yes		9 (22.0)	
RR criteria?	Yes		7 (17.0)	
BP criteria?	Yes		22 (53.6)	
AVPU criteria?	Yes		4 (9.7)	

Did the frequency of observations increase when MET was reached? E.g Repeat...	Yes No		26/45 19/45	

Appendix 2: Economic Analyses

Table 6: Observed training cost (6 months) for Web-based training group

Cost Type	Cost ($)	# of Participant
Filming	$8,169	
Software	$24,800	
Project Management	$19,515	
CGHS Trainee	$1,362	25
Bendigo Trainee	$1,726	30
Trainer	$942	
Total	**$56,514**	55 trainees
Mean 6 month Cost of Web-based training	$1027.53	

Table 7: Observed training cost (6 months) for Face2Face training group

Cost Type	Cost ($)	# of Participant
Project Management Costs	$15,910	
Berwick Trainee Costs	$2,804	34
LRH Trainee Costs	$2,906	40
Trainer Costs	$14,656	
Total Costs	**$36,276**	74 trainees
Mean 6 month Cost of Face2Face training	$490.22	

Table 8: Summary of hospitalization cost and associated factors for Web-based group

		# of cases	Mean	Minimum	Lower quartile	Median	Upper quartile	Maximum	STD
Overall (N=736)	Cost ($)	655	**$6492.8**	326.8	1709.8	3310.7	7608.6	89614.6	9242.9
	LOS (days)	736	**6.31**	1.0	2.0	4.0	8.0	45.0	5.9
	ICU (days)	9	**5.11**	2.0	4.0	4.0	8.0	8.0	2.32
	Weight SD[1]	302	0.29	0.11	0.20	0.23	0.32	0.75	0.15
	Weight OD[1]	128	0.35	0.23	0.33	0.35	0.38	0.76	0.07
	Weight MD[1]	225	1.71	0.06	0.76	1.32	1.84	10.57	1.77
Before training (N=383)	Cost ($)	339	**$6040.3**	326.8	1709.8	3065.9	6835.5	89614.6	9589.4
	LOS (days)	383	**5.85**	1.0	2.0	4.0	7.0	32.0	5.39
	ICU (days)	4	**5.25**	2.0	2.50	5.50	8.0	8.0	3.20
	Weight SD[1]	166	0.29	0.14	0.20	0.23	0.32	0.75	0.15
	Weight OD[1]	70	0.35	0.23	0.33	0.34	0.38	0.76	0.08
	Weight MD[1]	103	1.85	0.06	0.64	1.35	1.94	10.57	2.04
After training (N=352)	Cost ($)	315	**$6928.4**	350.0	1709.8	3776.6	8891.9	74883.3	8815.1
	LOS (days)	352	**6.77**	1.0	3.0	5.0	8.50	45.0	6.38
	ICU (days)	5	**5.0**	4.0	4.0	4.0	5.0	8.0	1.73
	Weight SD[1]	136	0.29	0.11	0.20	0.23	0.32	0.75	0.15
	Weight OD[1]	58	0.36	0.26	0.33	0.35	0.36	0.56	0.07
	Weight MD[1]	121	1.59	0.13	0.90	1.32	1.84	9.63	1.50

Cost = Total hospitalization cost (AUS $); LOS = Length of stay in ward (day); ICU = Length of stay in ICU (days); Weight SD = Resource Intensity Weight for same day procedure; Resource Intensity Weight OD = Weight for one day procedure; Resource Intensity Weight MD = Weight for multi-day procedure;
[1]Weis cost weights 2016-17. Available online: https://www2.health.vic.gov.au/about/publications/policiesandguidelines/wies23-cost-weights-2016-17

Table 9: Summary of hospitalization cost and associated factors for Face2Face group

		# of cases	Mean	Minimum	Lower quartile	Median	Upper quartile	Maximum	STD
Overall (N=827)	Cost ($)	809	**$5259.3**	194.3	1422.2	2643.5	5859.8	97235.4	7963.5
	LOS (days)	827	**4.94**	1.0	2.0	4.0	6.0	38.0	4.70
	ICU (days)	9	**4.33**	1.0	3.0	4.0	5.0	10.0	2.55
	Weight SD[1]	397	0.25	0.10	0.19	0.22	0.28	0.52	0.09
	Weight OD[1]	139	0.37	0.23	0.33	0.34	0.37	0.77	0.10
	Weight MD[1]	273	1.51	0.06	0.68	1.28	1.75	9.63	1.41
Before training (N=396)	Cost ($)	389	**$5280.6**	194.3	1396.2	2717.9	5692.4	84027.5	7797.5
	LOS (days)	396	**4.77**	1.0	2.0	4.0	6.0	38.0	4.43
	ICU (days)	4	**5.75**	3.0	4.0	5.0	7.5	10.0	2.99
	Weight SD[1]	191	0.25	0.10	0.19	0.21	0.28	0.52	0.09
	Weight OD[1]	69	0.37	0.23	0.33	0.34	0.38	0.76	0.11
	Weight MD[1]	129	1.52	0.06	0.63	1.19	1.84	9.63	1.39
After training (N=421)	Cost ($)	410	**$5302.1**	311.5	1440.3	2644.4	6787.1	97235.4	8206.1
	LOS (days)	421	**5.16**	1.0	2.0	4.0	6.0	36.0	4.98
	ICU (days)	5	**3.20**	1.0	2.0	4.0	4.0	5.0	1.64
	Weight SD[1]	202	0.25	0.10	0.19	0.22	0.28	0.49	0.09
	Weight OD[1]	69	0.37	0.23	0.33	0.35	0.36	0.77	0.10
	Weight MD[1]	139	1.54	0.14	0.78	1.28	1.71	8.04	1.44

Cost = Total hospitalization cost (AUS $); LOS = Length of stay in ward (day); ICU = Length of stay in ICU (days); Weight SD = Resource Intensity Weight for same day procedure; Weight OD = Resource Intensity Weight for one day procedure; Weight MD = Resource Intensity Weight for multi-day procedure

[1]Weis cost weights 2016-17. Available online: https://www2.health.vic.gov.au/about/publications/policiesandguidelines/wies23-cost-weights-2016-17

Table 10: Summary on hospitalization cost from bootstrap sample

Difference in Costs	Mean	2.5% and 97.5%percentile	Minimum	Median	Maximum	Standard Error
Face2Face	-$9.80	(-1068, 1114)	-2017.4	-25.8	1725.4	17.5
Web-based	$848.60	(-599, 2220)	-1270.4	843.9	2904.3	22.4
Web-based minus F2F	$858.30	(-973, 2648)	-1881.7	865.6	3458.5	28.6

Table 11: Summary on length of stay in ward from bootstrap sample

Difference in LOS	Mean	2.5% and 97.5%percentile	Minimum	Median	Maximum	Standard Error
Face2Face	0.39	(-0.23, 1.03)	-0.81	0.39	1.57	0.01
Web-based	0.90	(0.001, 1.74)	-0.28	0.89	2.30	0.01
Web-based minus F2F	0.51	(-0.62, 1.56)	-1.07	0.55	2.39	0.02

1. The first difference between before and after training is calculated as:
Difference for F2F = mean hospitalization cost after Face2Face training – mean hospitalization cost before Face2Face training
Difference for Web = mean hospitalization cost after web-based training – mean hospitalization cost before web-based training
2. The second difference between Face2Face and web-based training is calculated as:
Overall difference = Difference for web-based - Difference for Face2Face
3. Summary statistics for difference for Face2Face, web-based, and overall are reported based on 1000 bootstrap sample.
4. The 25th and 975th observations represent the 95% CI from each distribution of differences after sorted in ascending order.