Clinical Simulation in Nursing (2019) 32, 1-7

Clinical Simulation in Nursing

www.elsevier.com/locate/ecsn

Featured Article

Medication Administration Evaluation and Feedback Tool: Simulation Reliability Testing

Karen M. Davies, RN, BN[a,b,c,*], Ian D. Coombes, BPharm, MSc, PhD[c,d], Samantha Keogh, RN, PhD[e,f,g], Karen Hay, BVSc, MEpi, PhD[h], Cameron Hurst, BSc, PhD[h], Karen M. Whitfield, BPharm, MSc, PhD[c,d]

[a]*Department of Clinical Pharmacology, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia*
[b]*Faculty of Medicine, University of Queensland, Brisbane, Queensland, Australia*
[c]*School of Pharmacy, University of Queensland, Brisbane, Queensland, Australia*
[d]*Pharmacy Department, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia*
[e]*School of Nursing, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia*
[f]*Nursing and Midwifery Research Centre, Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia*
[g]*Alliance for Vascular Access Teaching and Research Group, Griffith University, Brisbane, Queensland, Australia*
[h]*Statistics Unit, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia*

KEYWORDS
intra-rater and inter-rater;
reliability;
medication administration;
nurses;
nurses performance evaluation;
feedback;
simulation

Abstract

Background: Incorrect medication administration risks errors and patient harm. The aim of this study was to test the reliability of the newly developed medication administration evaluation and feedback tool.

Methods: An observational, fully crossed design using recorded scenarios in a simulated environment was used to test reliability and agreement of an evaluation tool for nurses administering medications.

Results: Intrarater and inter-rater reliability observed agreement overall were 84% and 82% with Fleiss Kappa coefficient 0.72 and 0.68, respectively. Overall intrarater and inter-rater reliability of the tool rated as good.

Conclusions: The medication administration evaluation and feedback tool is reliable in a simulated environment.

Cite this article:
Davies, K. M., Coombes, I. D., Keogh, S., Hay, K., Hurst, C., & Whitfield, K. M. (2019, July). Medication administration evaluation and feedback tool: Simulation reliability testing. *Clinical Simulation in Nursing, 32*(C), 1-7. https://doi.org/10.1016/j.ecns.2019.03.010.

Background

Medication administration is one component of the medication management process of prescribing, dispensing, and administering. Nurses are the predominant administrators of medications. However, despite extensive education and training, studies show error rates between 10% and 25% persist, depending on the study design and outcome definitions (Berdot et al., 2013). A review of reported medication incidents in the United Kingdom over a 10-year

* Corresponding author: karen.davies@uqconnect.edu.au (K. M. Davies).

1876-1399/$ - see front matter
https://doi.org/10.1016/j.ecns.2019.03.010

period showed that nearly one-third of medication administration errors resulting in death were omitted medications (Härkänen, Vehviläinen-Julkunen, Murrells, Rafferty, & Franklin, 2018). The World Health Organization (WHO) in March 2017 launched the ''Medication Without Harm: WHO's Third Global Patient Safety Challenge'' with an aim to halve medication-related errors in 5 years (World Health Organisation, 2017). The WHO education and training resource strategy states that clinical supervision, role modelling, and feedback are core to improving medication safety (World Health Organisation, 2016).

Key Points

- Simulation can effectively test intra-rater and inter-rater reliability of a Medication Administration Evaluation Tool.
- The Medication Administration Evaluation and Feedback Tool is reliable.
- Nurses' medication administration practice is supported with self-reflection, performance evaluation, and feedback.

Levels of specific guidance on medication administration standards vary internationally. Initially, the United Kingdom Nursing and Midwifery Council Standard for Medicine Management (Nursing and Miwifery Council, 2010) provided details on steps for safe medication administration. However, these standards were withdrawn in 2018 because of a concern they could restrict more contemporary medication practice (Merrifield, 2017; Nursing and Midwifery Council, 2018). Other associations like the College of Nurses of Ontario have broad principles on medication practice in their Practice Standards for Medications (College of Nurses Ontario, 2017).

In Australia, Medication Safety standards are set out by the Australian Commission on Safety and Quality in Health Care (2017). The standard provides broad principles on medication administration with the expectation that local health services provide specific procedures on how to safely administer medications. Ensuring competent clinicians safely manage medicines is the intent of this standard. However, the standard does not specify how to determine if clinicians perform safely in practice.

This study was initiated because of the health care burden of medication errors and demonstrated evidence that observation and direct formative feedback showed a statistically significant improvement to the medication error rate (Davies, Mitchell, & Coombes, 2015a; Davies et al., 2015b). In conducting these studies, it became apparent there was a lack of a suitable and validated tool to assess nurses' medication administration practice. This led to the Medication Administration Evaluation and Feedback Tool (MAEFT) design from an expert panel review of three existing tools identified in the literature (Goodstone & Goodstone, 2013; Hemingway, Baxter, Smith, Burgess-Dawson, & Dewhirst, 2011; O'Brien, 2015). The MAEFT incorporates self-evaluation, peer observation, and feedback and is designed to develop an agreed performance development plan for continuing improvement of practice (Davies, Coombes, Keogh, & Whitfield, 2019) See Appendix A. As a newly developed tool, the designed MAEFT required testing for reliability and agreement.

Aim

The aim of the study was to assess the reliability of the MAEFT.

Objectives

The objectives of the study were to test the developed MAEFT for intra-rater and inter-rater reliability, agreement, and accuracy by nurse educators in a simulated clinical environment using digitally recorded scenarios.

Theoretical Framework

This study is based on the theory of andragogy, the Adult Learning Principles as described by Knowles (Knowles, 1950). The theory is based on the process of how adults learn using problem-based principles that are collaborative between the learner and the teacher. Principles include self-motivation, previous experience, relevance, practical, goal-orientated learning designed with mutual respect (Kaufman, 2003). As part of using this seminal theory, observation of clinical performance is widely used as a quality assessment strategy in health care to evaluate compliance with guidelines and deviation from best practice, potentially jeopardising patient safety (Colquhoun et al., 2013; Yanes et al., 2015).

The newly designed MAEFT incorporates self-evaluation, observation, and observer feedback based on safe practice steps to administer medications. How feedback is performed and received is critical to the benefit on learner performance. It is agreed by educational theorists that the most effective learners self-regulate and both internal and external feedback are integral in determining self-regulated learning (Butler & Winne, 1995; Sitzmann & Ely, 2011). Both nurses and their supervisors found that evaluation tools identifying specific criteria for self-assessment and feedback supported their learning in a clinical environment.

(Calleja, Harvey, Fox, & Carmichael, 2016). Feedback can be powerful but, depending on the type of feedback, the outcome can be positive or negative (Hattie & Timperley, 2007). For feedback to be effective, it must be observable, objective, respectful, and timely to support the learner to improve performance (Johnson et al., 2016).

Methods

Study Design/Setting

The study was an observational fully crossed design to test reliability and agreement of the MAEFT. That is, the same set of multiple observers rated all subjects (Hallgren, 2012). The modality used was a simulation-based experience (SBE) with a backstory of scripted digitally recorded scenarios of a nurse administering medications to a patient in a purpose designed, clinical skills laboratory, simulated environment conducted in November 2017. The case scenarios, medications, medication requirements, and orders were realistic to ensure conceptual fidelity. The setting was a simulated clinical ward bed environment with an adjoining medication preparation room fitted with a controlled drug safe to enhance physical/environmental fidelity (INACSL Standards Committee, 2016). The study was granted low-risk ethical approval by the Human Research Ethics Committee in August 2017, reference HREC/17/QRBW/402.

To help determine intra-rater and inter-rater reliability of the MAEFT, there were the following:

- Eight SBEs digitally recorded of a simulation nurse (SN) administering medications.
- Three simulation nurse educators (SNE) independently viewed the digitally recorded SBE scenarios in December 2017 and evaluated the nurse's medication administration practice using the MAEFT.
- The same three SNEs independently viewed the eight digitally recorded SBEs again seven days later using the MAEFT.

Simulated Patients/Nurses

The simulated patients (SPs) were role played by volunteer hospital staff, a health professional, and administration officer. The SNs administering medications were volunteer nurses, a medication safety intensive care nurse specialist, and a nurse academic. The volunteer health professional also fulfilled the role of the simulated medical officer. The SNE observers testing the MAEFT were nurse educators from the same tertiary metropolitan hospital.

The SNEs were all female and ranged from 30 to 70 years old. They were all registered nurses (RNs) with Bachelor qualifications with two having a Master's in Nursing Education. All had been practicing as RNs for between 10 and 30 years and in NE roles between 5 and 20 years. All SNE observers had experience with use of simulated environments for education and in observation and feedback of nursing practice.

Prebriefing

The researcher facilitated a prebriefing orientation of all participants. The two SNs were orientated to the SBE scenario scripts in a two-hour meeting the week before the scheduled scenario recordings. The SNs were also orientated to the clinical skills ward and medication room equipment and environment. The SPs were orientated to the medication script scenarios and clinical skills ward environment which included a moulage of patient gowns, broken spectacles, intravenous line arm adaptor, intravenous infusion pump, subcutaneous injection abdomen pad props, and urine drainage catheter bag to increase fidelity of the SBE.

The three SNE observers of the scenarios were orientated to the use of the MAEFT in a one-hour PowerPoint overview prepared by the researcher. The session was conducted the week before the first observation session and included the background, rationale and detailed design of the tool contents, and how to use it. However, as the tool was new, the training did not include practical demonstration and use of the tool. Therefore, there was no required minimum level of inter-rater reliability achievement before participation and observation of the nurse scenario recordings using the MAEFT. Evaluation of the training, however, was rated as practical and useful and the process experience as positive.

Theoretical Rationale for the Design of the Intervention

The theoretical framework for use of the simulation design was based on ''The National League for Nursing (NLN) Jeffries Simulation Theory'' (Jeffries, Rodgers, & Adamson, 2015). The theory context looks at factors such as the purpose of the simulation and incorporates background, design, simulation experience, facilitator and educational strategies, participants, and outcomes. SBE learning is founded on the principles of the Adult Learning Theory (Wang, 2011). For the context of this study, simulation was used for evaluation purposes. The SBE allowed fixed scenarios that were digitally recorded to be reproducible to enable intra-rater and inter-rater reliability assessment of the MAEFT (Walshe, O'Brien, Hartigan, Murphy, & Graham, 2014).

Simulation Scenarios

The researcher wrote, facilitated, and directed the eight digitally recorded SBE scenarios with intentionally scripted medication errors. There were preprepared packs for each scenario with medication charts for preadministration and postadministration documentation and labelled medication packages. Additional administration equipment including syringes, bags, IV lines, and infusion pumps with medication library dosing safety software, computer programmes with medication resource information, and vital sign monitoring equipment were made available. The SBE scenarios were digitally recorded and edited in postproduction by a professional film maker under the direction of the researcher. There

were two patients with multiple episodes of administration for each. See Appendix B for details of the scenarios.

Data Collection

The MAEFT comprises 22 questions to evaluate the nurses' medication administration practice. The MAEFT was used by three SNE raters on two occasions to assess eight digitally recorded SBE scenarios. Demographic data for the SNE raters collected included gender, age, academic qualifications, how many years they had been practicing as an RN, and how many years they had been practicing in their nurse educator role.

Data collected using the MAEFT were entered into the web-based Vanderbilt University—designed software Research Electronic Data Capture, version 8.5.0 (Harris et al., 2009). All data entered were de-identified.

Statistical Analysis

The categorical rating scale in the MAEFT was "yes," "no," or "not applicable," which was coded on a nominal scale. Each SNE observer viewed eight SBE with 22 questions equalling 176 ratings per SNE rater. Each SNE rater repeated the process seven days later, taking the total number of ratings to 352. The consistency of agreement for each SNE rater between time points (intra-rater) and between SNE raters at each time point (inter-rater) was determined using Fleiss' Kappa coefficient (Fleiss, Levin, & Paik, 2004), an extension of the Kappa statistic used for nominal data, allowing assessment by multiple raters. Fleiss' kappa coefficient also corrects for agreement between ratings because of chance. A fully crossed two-way random effects design was used treating variation due to rater and scenario as random (e.g., a different random set of raters would not follow the particular, or fixed, difference observed with these particular raters). The evaluation criteria guide provided by Polit & Beck, 2006 was used to gauge agreement. Specifically, kappa agreement is poor if $\kappa < 0.40$, fair when κ is 0.40 to 0.59, good when κ is 0.60 to 0.74, and excellent for $\kappa > 0.74$. Due to the fully crossed design with multiple raters, an average percentage agreement was also calculated as proposed by biostatisticians Davies and Fleiss (1982). Subgroup analysis for each individual question was also calculated for agreement using Fleiss' Kappa coefficient. In some cases, it was not possible to calculate Fleiss' Kappa coefficient as there was perfect agreement. In these cases, the average percentage agreement was presented. All statistical analysis was conducted with assistance from two statisticians using the Stata (v15, StataCorp, College Station, 2017) and R statistical package (v3.4.4, R Core Team, 2018) with the R package irr used to calculate Fleiss' Kappa coefficient (Davies & Fleiss, 1982; Gamer, Lemon, Fellows, & Singh, 2012).

Results and Discussion

Measures of Reliability and Agreement

The individual SNE results rating the consistency on how often they agreed when using the MAEFT to rate the SNs administering medications in each of the eight SBE digitally recorded scenarios are shown in Table 1 (intra-rater) and Table 2 (inter-rater).

The SNE observer intra-rater agreement percentage and overall comparison ratings for both time points ranged between 81.25% and 84.28%. When agreement by chance is accounted for, the overall Fleiss' Kappa coefficient statistic for intra-rater reliability was 0.72, and for inter-rater reliability, it was 0.68. The observed agreement of 84.28% and Fleiss' Kappa coefficient of 0.72 indicates good agreement therefore good intra-rater reliability.

The individual results were supported by a combined overall comparison of raters with an average percentage agreement for each of the 22 criteria/questions of 81.91% and a Fleiss' Kappa coefficient of 0.68. Inter-rater results evaluated as good agreement and reliability. The accuracy of the MAEFT was calculated by the average percentage of the correct answer. The combined overall comparison of raters with the correct answer showed similar results with an accuracy average of 80.87%.

When scrutinising the inter-rater reliability of MAEFT at time one, time two, and both times combined (Table 2), there was some variability in the level of agreement, although all Fleiss' Kappa coefficient values were within the good range. Inter-rater reliability seemed to be systematically lower at time two, relative to time one. This was investigated further with the average percentage agreement per question. Question two, 14, 15, and 20 decreased by more than 10%

Table 1 Intra-rater Reliability Measured Using the Fleiss' Kappa Coefficient Statistic

Comparison	Observed Agreement (%)	Expected Agreement (%)	Fleiss' Kappa Coefficient	95% CI	Evaluation
Overall comparisons of ratings between time 1 and 2	84.28	43.13	0.7236	(0.66,0.79)	Good
Comparison for observer 1	84.09	42.69	0.7224	(0.63,0.82)	Good
Comparison for observer 2	87.50	41.75	0.7854	(0.70,0.87)	Excellent
Comparison for observer 3	81.25	45.25	0.6575	(0.55,0.76)	Good

Table 2 Inter-Rater Reliability and Accuracy (% Correct) Measured Using the Fleiss' Kappa Coefficient Statistic

Comparison	Outcome	Av % Agreement	Fleiss' Kappa Coefficient	95% CI	Evaluation	Av % Correct
Comparing raters at time 1	1		0.74		Good	79.55
	2		0.71		Good	83.52
	3		0.71		Good	84.09
	Combined	84.47	0.73	(0.65,0.80)	Good	82.39
Comparing raters at time 2	1		0.65		Good	81.82
	2		0.64		Good	80.11
	3		0.61		Good	76.14
	Combined	79.35	0.64	(0.56,0.72)	Good	79.36
Overall comparison of raters	1		0.70		Good	80.68
	2		0.68		Good	81.82
	3		0.66		Good	80.11
	Combined	81.91	0.68	(0.63,0.74)	Good	80.87

Fleiss' Kappa coefficient agreement evaluation criteria: poor = $\kappa < 0.40$; fair = κ 0.40 to 0.59; good = κ 0.60 to 0.74; excellent = $\kappa > 0.74$.

of average percentage agreement at time two as seen in Table 3.

Significance of Results

The Fleiss' Kappa coefficient measured evaluation criteria as good for intra-rater and inter-rater reliability of SNE observers using the MAEFT. The significance of these results is that they demonstrate that the newly designed MAEFT is reliable when used by multiple observers to observe different SN/SP scenarios where there is a fixed SBE. As all SNs were observed by the same SNE raters at both time points, the inter-rater reliability SBE outcomes of the Fleiss' Kappa coefficient results could be generalised with use of the MAEFT. However, further studies testing the MAEFT in a clinical environment are warranted.

Table 3 Question Average Percentage Agreement

Question	Av % Agreement T1	Av % Agreement T2	% Difference	Av % Agreement Both T
1	91.66	91.66	0.00	91.66
2	100.00	66.66	33.34	83.33
3	100.00	91.66	8.34	95.83
4	100.00	100.00	0.00	100.00
5	83.33	91.66	−8.33	87.50
6	66.66	75.00	−8.34	70.83
7	91.66	91.66	0.00	91.66
8	100.00	91.66	8.34	95.83
9	100.00	100.00	0.00	100.00
10	91.66	91.66	0.00	91.66
11	75.00	66.66	8.34	70.83
12	58.33	58.33	0.00	58.33
13	91.66	91.66	0.00	91.66
14	66.66	50.00	16.66	58.33
15	91.66	62.50	29.16	77.08
16	91.66	91.66	0.00	91.66
17	75.00	75.00	0.00	75.00
18	75.00	75.00	0.00	75.00
19	83.33	75.00	8.33	79.17
20	75.00	62.50	12.50	68.75
21	66.66	62.50	4.16	64.58
22	83.33	83.33	0.00	83.33
Mean	84.47	79.35	5.11	81.91
Range	58-100	50-100	0-33	58-100
SD	12.93	14.85	−1.92	12.93
SE	2.82	3.24	−0.42	2.82
Lower	78.60	72.61	5.98	76.04
Upper	90.34	86.10	4.24	87.78

Despite the good rating of the MAEFT, it was noted there was variability in the Fleiss' Kappa coefficient inter-rater results between time points with time two lower than time one. This was broken down into the average percentage agreement where some questions, example question two, had three patient identifiers to be checked on the patient identification band and the medication chart. If the identification band was checked and not cross referenced with the medication chart, then all criteria were not met, and the answer would be ''no'' as it was not checked completely. Question 14 addressed aseptic technique when preparing and administering the medication. Although this required appropriate hand hygiene as well question 13 specifically addresses this to differentiate between equipment asepsis and hand hygiene. Question 20 is about engaging the patient to determine their knowledge of the medications received. If the engagement used closed questions and not open, it may have been interpreted as poor engagement and therefore ''no'' depending on the observer. An example of this is the nurse stating, ''here is your antihypertensive for your blood pressure'' rather than ''here is your perindopril do you know what you take that for?''

Possible reasons for inter-rater results being lower at time two may be due to changes in interpretation and understanding of criteria in the tool. It may also be due to variation in the speed observers grasped the use of the tool. Observers confidence with observing and evaluating the scenarios could improve with practice and familiarity. This is not unexpected for this first stage of assessing the reliability of the tool. This could be addressed by optimised education of the observer before participating in observation of nurses administering medications using the MAEFT. An education strategy is planned, and this would be an essential component of the use of the MAEFT to ensure observer competence.

Limitations

There are some limitations for this study. The SNE received limited training with no practical demonstration and use of the MAEFT. The SNE did not have to achieve a required minimum level of intra-rater reliability before participation and observation of the SN digitally recorded scenarios. The achieved intra-rater reliability could have been lower as a result. To address this in future, the simulation scenario recordings will be used for training using the MAEFT. A minimum required level of achievement before clinical use will also be considered. As the MAEFT was tested in a simulated environment only, results may not translate to a practical/clinical environment. So further testing of the tool in the clinical environment is warranted.

Conclusions

This study of the newly developed MAEFT has demonstrated reliability and accuracy in a simulated environment. Further studies of the MAEFT to determine clinical reliability and generalisability have been conducted in a clinical environment and results are undergoing analysis. A follow-up postpilot study has also been conducted to assess whether using the MAEFT makes a difference to medication administration practice. Results are yet to be analysed. Further research developing an education plan to use the MAEFT is required to ensure consistency and maximum benefit as a valid and reliable assessment.

Acknowledgment

The authors would like to acknowledge the additional actors Jennifer Philp and Kahlya Simpson for the simulation and Luke Mayze for his expertise in film directing.

Funding: This work was supported and funded by the Queensland Health, Metro North Hospital and Health Services, Royal Brisbane & Women's Hospital, Research Postgraduate Scholarships 2017 and 2018 for the first author PhD candidate through the University of Queensland.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecns.2019.03.010.

References

Australian Commission on Safety and Quality in Health Care. (2017). *National Safety and Quality Health Service Standards* (2nd ed.). Sydney: ACSQHC.

Berdot, S., Gillaizeau, F., Caruba, T., Prognon, P., Durieux, P., & Sabatier, B. (2013). Drug administration errors in hospital inpatients: a systematic review. *PLoS One*, *8*(6), e68856. https://doi.org/10.1371/journal.pone.0068856.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, *65*(3), 245.

Calleja, P., Harvey, T., Fox, A., & Carmichael, M. (2016). Feedback and clinical practice improvement: a tool to assist workplace supervisors and students. *Nurse Education in Practice*, *17*(Suppl C), 167-173. https://doi.org/10.1016/j.nepr.2015.11.009.

College of Nurses Ontario. (2017). *Medication Practice Standard*. Retrieved from http://www.cno.org/en/learn-about-standards-guidelines/standards-and-guidelines/. (Accessed 23 June 2018).

Colquhoun, H. L., Brehaut, J. C., Sales, A., Ivers, N., Grimshaw, J., Michie, S., Carroll, K., Chalifoux, M., & Eva, K. W. (2013). A systematic review of the use of theory in randomized controlled trials of audit and feedback. *Implementation Science: IS*, *8*, 66. https://doi.org/10.1186/1748-5908-8-66.

Davies, K., Coombes, I., Keogh, S., & Whitfield, K. (2019). Medication administration evaluation tool design: an expert panel review. *Collegian*, *26*(1), 118-124. https://doi.org/10.1016/j.colegn.2018.05.001.

Davies, M., & Fleiss, J. L. (1982). Measuring agreement for multinomial data. *Biometrics*, *38*(4), 1047-1051.

Davies, K., Mitchell, C., & Coombes, I. (2015a). The role of observation and feedback in enhancing performance with medication administration. *Journal of Law and Medicine*, *23*(2), 316.

Davies, K., Norris, L., O'Brien, E., Buksh, F., Rogers, K., Kubler, P., Donovan, P., & Coombes, I. (2015b). Internal Medicine Services (IMS) medication administration observation & feedback. In *Paper Presented at the Royal Brisbane and Women's Hospital Symposium, Brisbane*.

Fleiss, J. L., Levin, B., & Paik, M. C. (2004). *Statistical Methods for Rates and Proportions* (3rd ed.). Hoboken: Wiley.

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). Irr: various coefficients of interrater reliability and agreement. R package version 0.84. Retrieved from https://CRAN.R-project.org/package=irr. (Accessed 26 April 2018).

Goodstone, L., & Goodstone, M. S. (2013). Use of simulation to develop a medication administration safety assessment tool. *Clinical Simulation in Nursing*, *9*(12), e609-e615. https://doi.org/10.1016/j.ecns.2013.04.017.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23. https://doi.org/10.20982/tqmp.08.1.p023.

Härkänen, M., Vehviläinen-Julkunen, K., Murrells, T., Rafferty, A. M., & Franklin, B. D. (2018). Medication administration errors and mortality: Incidents reported in England and Wales between 2007 2016. *Research in Social and Administrative Pharmacy*. https://doi.org/10.1016/j.sapharm.2018.11.010.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377-381. https://doi.org/10.1016/j.jbi.2008.08.010.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487.

Hemingway, S., Baxter, H., Smith, G., Burgess-Dawson, R., & Dewhirst, K. (2011). Collaboratively planning for medicines administration competency: A survey evaluation. *Journal of Nursing Management*, *19*(3), 366-376. https://doi.org/10.1111/j.1365-2834.2011.01245.x.

INACSL Standards Committee. (2016). INACSL standards of best practice: SimulationSM simulation glossary. *Clinical Simulation in Nursing*, *12*, S39-S47. https://doi.org/10.1016/j.ecns.2016.09.012.

Jeffries, R. P., Rodgers, R. B., & Adamson, R. K. (2015). NLN Jeffries simulation theory: brief narrative description. *Nursing Education Perspective*, *36*(5), 292-293.

Johnson, C., Keating, J., Boud, D., Dalton, M., Kiegaldie, D., McKenzie, W., Nestel, D., Palermo, C., & Molloy, E. (2016). Identifying educator behaviours for high quality verbal feedback in health professions education: literature review and expert refinement. *BMC Medical Education*, *16*, 96. https://doi.org/10.1186/s12909-016-0613-5.

Kaufman, D. M. (2003). Applying educational theory in practice. *BMJ*, *326*(7382), 213-216. https://doi.org/10.1136/bmj.326.7382.213.

Knowles, M. S. (1950). *Informal Adult Education: A Guide for Administrators, Leaders, and Teachers*. New York: Association Press.

Merrifield, N. (2017). RCN concerns over proposed changes to mentor training and medicines management. *Nursing Times*. Retrieved from https://www.nursingtimes.net/news/education/rcn-concern-at-plans-for-mentorship-and-meds-management/7021297.article. (Accessed 10 October 2017).

Nursing and Midwifery Council. (2010). *Standards for Medicines Management*. UK. (Accessed 5 April 2018).

Nursing and Midwifery Council. (2018). New NMC standards shape the future of nursing for next generation [Press release]. Retrieved from https://www.nmc.org.uk/news/press-releases/new-nmc-standards-shape-the-future-of-nursing-for-next-generation/. (Accessed 3 December 2018).

O'Brien, E. (2015). Medication resource package. RBWH intranet: royal brisbane and women's hospital (RBWH). Retrieved from http://hi.bns.health.qld.gov.au/nursing/education.htm#workforce_development. (Accessed 10 March 2016).

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489-497. https://doi.org/10.1002/nur.20147.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychological Bulletin*, *137*(3), 421-442. https://doi.org/10.1037/a0022777.

Walshe, N., O'Brien, S., Hartigan, I., Murphy, S., & Graham, R. (2014). Simulation performance evaluation: Inter-rater reliability of the DARE2-patient safety rubric. *Clinical Simulation in Nursing*, *10*(9), 446-454. https://doi.org/10.1016/j.ecns.2014.06.005.

Wang, E. E. (2011). Simulation and adult learning. *Disease-a-Month*, *57*(11), 664-678. https://doi.org/10.1016/j.disamonth.2011.08.017.

World Health Organisation. (2016). *Education and Training: Technical Series on Safer Primary Care. (Licence: CC BY-NC-SA 3.0 IGO.)*. Geneva World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/10665/252271/1/9789241511605-eng.pdf?ua=1. (Accessed 15 April 2017).

World Health Organisation. (2017). Medication without harm: WHO's third global patient safety challenge. Retrieved from http://www.who.int/patientsafety/medication-safety/en/. (Accessed 15 April 2017).

Yanes, A. F., McElroy, L. M., Abecassis, Z. A., Holl, J., Woods, D., & Ladner, D. P. (2015). Observation for assessment of clinician performance: a narrative review. *BMJ Quality & Safety*, *25*, 46-55. https://doi.org/10.1136/bmjqs-2015-004171.