International Journal of Nursing Studies 52 (2015) 1231–1242

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

Review

12 h shifts and rates of error among nurses: A systematic review

Jill Clendon [a,b,*], Veronique Gibbons [c,d]

[a] Victoria University, Wellington, New Zealand
[b] New Zealand Nurses Organisation, PO Box 1195, Nelson 7040, New Zealand
[c] Waikato District Health Board, Hamilton, New Zealand
[d] School of Nursing, University of Auckland, New Zealand

ARTICLE INFO

Article history:
Received 20 October 2014
Received in revised form 8 March 2015
Accepted 17 March 2015

Keywords:
Error
Hospitals
Nurses
12 h shifts

ABSTRACT

Objective: To determine the effect of working 12 h or more on a single shift in an acute care hospital setting compared with working less than 12 h on rates of error among nurses.
Design: Systematic review.
Method: A three-step search strategy was utilised. An initial search of Cochrane, the Joanna Briggs Institute (JBI), MEDLINE and CINAHL was undertaken. A second search using all identified keywords and index terms was then undertaken across all included databases (Embase, Current contents, Proquest Nursing and Allied Health Source, Proquest Theses and Dissertations, Dissertation Abstracts International). Thirdly, reference lists of identified reports and articles were searched for additional studies. Studies published in English before August 2014 were included.
Findings: Following review of title and abstract of 5429 publications, 26 studies were identified as meeting the inclusion criteria and selected for full retrieval and assessment for methodological quality. Of these, 13 were of sufficient quality to be included for review. Six studies reported higher rates of error for nurses working greater than 12 h on a single shift, four reported higher rates of error on shifts of up to 8 h, and three reported no difference. The six studies reporting significant rises in error rates among nurses working 12 h or more on a single shift comprised 89% of the total sample size ($N = 60{,}780$ with the total sample size $N = 67{,}967$).
Conclusion: The risk of making an error appears higher among nurses working 12 h or longer on a single shift in acute care hospitals. Hospitals and units currently operating 12 h shift systems should review this scheduling practice due to the potential negative impact on patient outcomes. Further research is required to consider factors that may mitigate the risk of error where 12 h shifts are scheduled and this cannot be changed.

What is already known about the topic?

- 12 h shifts are a common shift pattern for nurses.
- Long work hours result in adverse nurse outcomes.
- The relationship between long work hours and adverse patient outcomes has been less clear.

What this paper adds

- The risk of making an error rates appears higher among nurses working 12 h or longer on a single shift in acute care hospitals.

* Corresponding author at: New Zealand Nurses Organisation, PO Box 1195, Nelson 7040, New Zealand. Tel.: +64 3 5463941.
E-mail address: jillc@nzno.org.nz (J. Clendon).

http://dx.doi.org/10.1016/j.ijnurstu.2015.03.011
0020-7489/

- Hospitals and units currently operating 12 h shift systems should review this scheduling practice due to the adverse impact on patient and nurse outcomes.

1. Introduction/background

Work schedules should be optimised for both employees and patients. The introduction of 12-h shifts into rostering/scheduling systems has been one approach implemented in workplaces with the intention of improving the flexibility of work hours for nurses. The approach has proven popular, with many proponents citing good quality time off work, ease of travel to work, improved relationships with patients, and better family time as benefits (Estryn-Béhar et al., 2012; O'Connor, 2011; Richardson et al., 2007). However, there is significant debate in the literature regarding the disadvantages of 12 h or longer shifts with some authors claiming extended shifts cause increased fatigue, greater risk of errors, greater risk of injury to self, and negative physiological outcomes (Chen et al., 2011; Estryn-Béhar et al., 2012; Geiger-Brown and Trinkoff, 2010; Rogers et al., 2004; Scott et al., 2006), others claiming no difference in patient outcomes (Stone et al., 2006), and yet others attributing greater risk of error to poor scheduling practices rather than length of shift (Estryn-Béhar et al., 2012). Error (see below for further definition) can lead to poor patient outcomes, therefore it is essential there is a good understanding of associated links between shift length and error to ensure optimal patient outcomes can be achieved. This paper presents the results of a systematic review undertaken to examine the relationship between 12 h shifts and error.

2. Literature review

Previous reviews have explored links between shift length and patient outcomes with results described as inconclusive (Bae and Fabry, 2014; Estabrooks et al., 2009). Bae and Fabry examined the relationship between nurse work hours/overtime and nurse and patient outcomes. They determined there is strong evidence supporting a positive relationship between working long hours and adverse nurse outcomes; however, their findings regarding shift length were not definitive – while they note that working more than 8.5 h per shift is significantly related to adverse patient outcomes, they also state more evidence is required to draw a definitive conclusion on the links between long work hours and adverse patient outcomes. Bae and Fabry do not explicitly define or examine error, nor do they focus specifically on 12 h versus 8-h shifts. Estabrook et al. examined the effect of shift length on quality of patient care and/or health provider outcomes but were unable to determine a clear finding due to the poor methodological integrity of the studies included in the review. Their review did include a small number of articles that specifically examined error but this was not their main focus. A further review of interest was that undertaken by Wagstaff and Sigstad (2011) who examined the safety implications of shift and night work and long working hours. Their findings suggest that work periods greater than 8 h carry an increased risk of accidents that accumulates with every further hour worked, but again, the review was not specific to studies on nurses or nursing. Ker et al. (2010) considered the effect of caffeine in preventing injuries and errors in shift workers and found when compared with no intervention, caffeine improved cognitive performance and reduced the number of errors in shift workers. Other reviews have examined the prevention of clinical error (Hodgkinson et al., 2006), the effect of flexible working conditions on employee health and wellbeing (Joyce et al., 2010), and hospital nurse staffing models and patient and staff-related outcomes (Butler et al., 2010). None of these reviews looked specifically at 12-h shifts.

There is wide variation among the findings of existing research studies that have examined 12-h shifts as noted above and the lack of any systematic examination of the evidence of associations between 12-h shifts, nurses and error provide justification for undertaking this review. It is essential for hospital managers and nurses to understand the ramifications of selected shift scheduling systems to ensure optimal patient outcomes.

3. Aim

The aim of this systematic review was to determine the effect of working 12 h or more on a single shift in an acute care hospital setting compared with working less than 12 h on rates of error among nurses. The question posed was: what is the effect of working 12 h or more on a single shift in an acute care hospital setting compared with working less than 12 h on rates of error among nurses?

4. Method

The review is based on the Joanna Briggs Institute systematic review process. This approach covers search strategy, inclusion/exclusion criteria, methodological quality, results and data synthesis.

4.1. Search strategy

To avoid duplication, an extensive search of the Cochrane library and the Joanna Briggs Institute was undertaken to ensure there was no existing systematic review on this topic nor any under development. The search strategy aimed to find both published and unpublished studies. A three-step search strategy was utilised. An initial limited search of MEDLINE and CINAHL was completed followed by examination of the key words and phrases contained in the title and abstract, and of the index terms used to describe the study. A second search using all identified keywords and index terms was then undertaken across all included databases. Thirdly, the reference lists of all identified reports and articles were searched for additional studies.

The databases searched included:

- CINAHL
- MEDLINE
- Embase
- Current contents
- Proquest Nursing and Allied Health Source

The search for unpublished studies included:

- Proquest Theses and Dissertations
- Dissertation Abstracts International

Initial search terms for all databases included:

12 h shifts OR shift work OR work pattern
AND
Nurs*
AND
clinical error OR practice error OR error OR medication error OR needle stick injury OR procedural error OR transcription error OR charting error OR incidents OR incident reporting OR hospital incidents OR safety OR patient safety OR safe practice OR safety events OR administration error OR event reporting OR failure OR safety OR lack of attentiveness OR lack of agency OR inappropriate judgement OR missed orders OR lack of intervention OR documentation error OR lack of prevention OR nursing error OR accident OR patient safety
AND
hospital OR acute care OR tertiary setting OR secondary setting

4.2. Inclusion/exclusion criteria

Studies published in English and before August 2014 were considered for inclusion in this review. The review included any studies with a quantitative, observational design that were studies of nurses (registered nurses (RNs), enrolled nurses (ENs), licensed practical nurses (LPNs), and licensed vocational nurses (LVNs) where relevant) working in acute care hospital settings, examined 12-h or longer shift patterns compared to studies examining shift patterns of less than 12 h, and considered rates of error as defined by Benner et al. (2002) (see below) among nurses. The study excluded studies of nursing students who are generally required to work under the supervision of a registered nurse, and other health and medical professionals whose shift patterns frequently differ from those of nurses. For the purposes of this review, the definition of error is based on Benner et al.'s taxonomy of nursing error (Benner et al., 2002). The taxonomy identifies eight categories of error including:

- lack of attentiveness (e.g. missing predictable complications such as post operative hemorrhage or poor monitoring of IV medications);
- lack of agency/fiduciary concern (e.g. failure to advocate for a patient's best interests, failure to question an inappropriate order, or a breach of confidentiality);
- inappropriate judgement (e.g. inadequate assessment, inability to recognise the implications of clinical signs and symptoms, reliance on prior convention or practice, unwarranted or faulty intervention);
- medication error (e.g. wrong drug, wrong route, wrong time, wrong patient, etc.);
- missed or mistaken doctor/health care provider orders (e.g. carrying out inappropriate orders, or mistaking orders resulting in erroneous intervention);
- lack of intervention on the patient's behalf (e.g. failure to follow up on signs or symptoms, lab results, etc.);
- documentation errors (e.g. charting procedures or medications before they were complete, failure to document patient observations);
- lack of prevention (failure to prevent threats to patient safety e.g. breach of infection control precautions or failure to prevent falls).

(Benner et al., 2002; Johnstone and Kanitsaki, 2006).

4.3. Search outcomes

Using the search terms, 5429 publications were identified using the initial search strategy. After review of the title and keywords, 5344 of these studies were rejected as not meeting the criteria for inclusion in the review (see above). These studies did not examine nurses or error or acute care. This left 86 studies as potentially relevant and full abstracts were reviewed. Following review of the abstract, 60 of these studies were then identified as not meeting the study's inclusion criteria i.e. did not specifically examine nursing, nurses, 12 h shifts or error; 26 studies were identified as broadly meeting the criteria for inclusion, retrieved in full and assessed for methodological quality as per the JBI critical appraisal process by both authors independently. Fig. 1 outlines the search strategy and outcomes.

4.4. Methodological quality

The 26 studies identified as broadly meeting the criteria for inclusion and retrieved in full were assessed using the JBI critical appraisal checklist for comparable cohort/case control studies or the JBI critical appraisal checklist for descriptive/case series studies (JBI, 2011) depending on the type of study. Of the 26 studies assessed for methodological quality, 13 were assessed as of sufficient methodological quality for inclusion – that is all met the JBI methodological criteria for having used reliable outcome measures and appropriate statistical analysis (questions eight and nine of the JBI critical appraisal tool). Excluded studies and the reasons for exclusion are found in Table 1. Tables 2a and 2b outline the quality assessment of the included studies. Table 3 details the included studies.

Where available, quantitative data was pooled in statistical meta-analysis using JBI-MAStARI. All results were subject to double data entry. Effect sizes expressed as odds ratio (for categorical data) and weighted mean differences (for continuous data) and their 95% confidence intervals were calculated for analysis. Heterogeneity was assessed statistically using the standard Chi-square and also explored using subgroup analyses based on the different study designs included in this review. Heterogeneity was identified by visual inspection of the forest plots, by using a standard χ^2 test and a significance level of $\alpha = 0.05$. Where $p < 0.05$ this is indicative of the studies being heterogeneous (JBI, 2011). Unfortunately the high degree of heterogeneity of the pooled data (heterogeneity Chi squared = 34.29, $p = 0.01$), meant we were unable to draw any substantial conclusions from the meta-analysis.

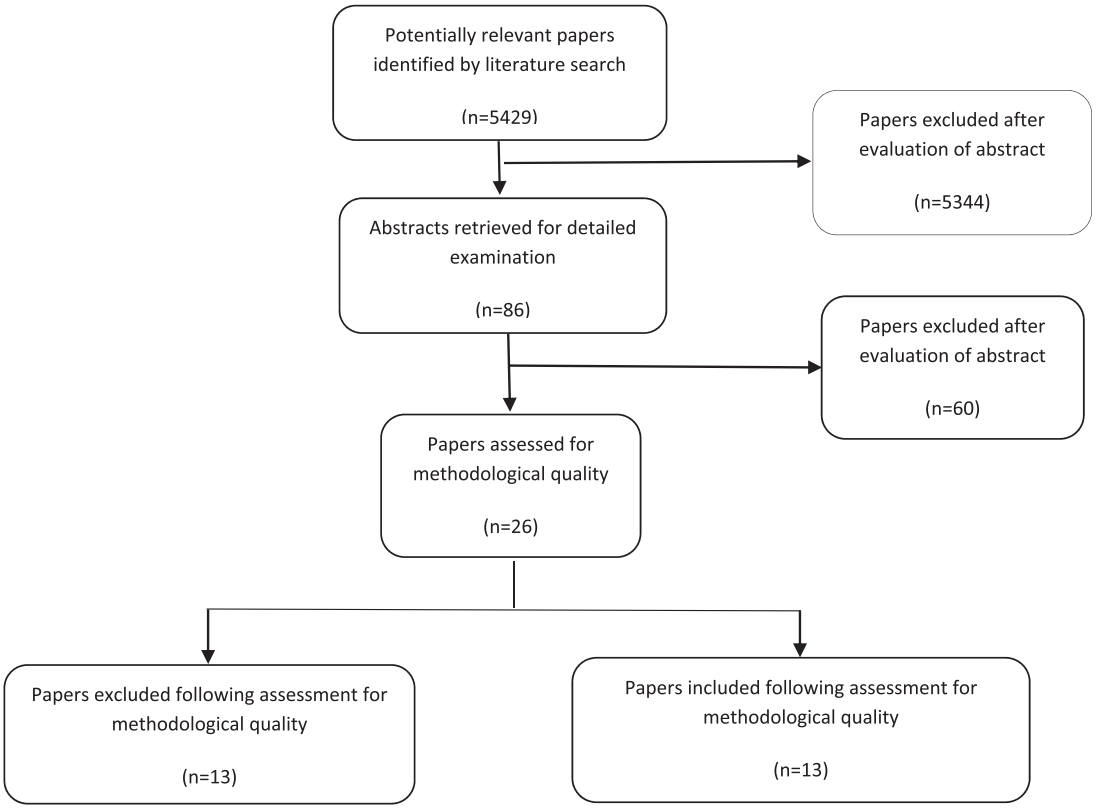

Fig. 1. Search strategy and outcome.

We have therefore presented the findings in narrative form including tables and figures to aid in data presentation where appropriate.

5. Results

5.1. Description of studies

Of the 13 studies included, two were from Japan (Arimura et al., 2010; Tanaka et al., 2010), one was from Australia (Dorrian et al., 2006), one was from Canada (Wilkins and Shields, 2008), one was from Europe (Griffiths et al., 2014), and eight were from the United States of America (Mills et al., 1983; Ritter, 1981; Rogers et al., 2004; Scott et al., 2006; Stone et al., 2006; Trinkoff et al., 2007; Witkoski Stimpfel et al., 2012, 2013). For the purposes of critical analysis using the JBI Critical Appraisal Checklist, the studies were identified into one of two categories with the majority being descriptive/case series studies and two were comparable cohort/case control studies (Mills et al., 1983; Trinkoff et al., 2007).

5.2. Samples

Studies used two different sampling techniques: convenience sampling (Arimura et al., 2010; Dorrian et al., 2006; Mills et al., 1983; Ritter, 1981; Stone et al., 2006; Tanaka et al., 2010; Witkoski Stimpfel et al., 2012, 2013) and random sampling (Griffiths et al., 2014; Rogers et al., 2004; Scott et al., 2006; Trinkoff et al., 2007; Wilkins and Shields, 2008). The number of participants ranged from 23 (Dorrian et al., 2006) to 33,659 (Griffiths et al., 2014). The average sample size was 5385 (median = 805).

5.3. Participants

The participants in each of the studies were similar in terms of characteristics. Most studies included RNs only, while others described participants as hospital nurses (Arimura et al., 2010; Stone et al., 2006), staff nurses (Witkoski Stimpfel et al., 2012), or included RNs, LPNs, and RPNs (Wilkins and Shields, 2008), or RNs and ENs (Dorrian et al., 2006). The average age of participants per study ranged from 28 years (Tanaka et al., 2010) to 44.8 years (Rogers et al., 2004). The majority of participants in all studies were female (>94%).

5.4. Settings

Study settings covered a broad range of hospital settings. These included medical/surgical (Griffiths et al., 2014), general hospital (Arimura et al., 2010; Dorrian et al., 2006; Rogers et al., 2004; Stone et al., 2006; Tanaka et al., 2010; Trinkoff et al., 2007; Wilkins and Shields, 2008; Witkoski Stimpfel et al., 2013), intensive care, coronary

Table 1
Excluded studies and reasons for exclusion.

Barker, L. M., & Nussbaum, M. A. (2011). Fatigue, performance and the work environment: A survey of registered nurses. *Journal of Advanced Nursing, 67*(6), 1370–1382. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Bloodworth, C., Lea, A., Lane, S., & Ginn, R. (2001). Challenging the myth of the 12-hour shift: A pilot evaluation. *Nursing Standard (Royal College of Nursing (Great Britain): 1987), 15*(29), 33–36. **Reason for exclusion:** Study considered outcomes from a change in shift patterns therefore not related to review objectives.
Estryn-Béhar, M., & Beatrice I.J.M.H. (2012). Effects of extended work shifts on employee fatigue, health, satisfaction, work/family balance, and patient safety. *Work, 41*, 4283–4290. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Fields, W. L., & Loveridge, C. (1988). Critical thinking and fatigue: How do nurses on 8- & 12-hour shifts compare? *Nursing Economic$, 6*(4), 189–191. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Jones, J. J., & Brown, R. M. (1986). A survey of the 12-hour nursing shift in 25 north Carolina hospitals. *Nursing Management, 17*(5), 27–28. **Reason for exclusion:** Survey was of hospitals not nurses, therefore did not meet review objectives.
Lea, A., & Bloodworth, C. (2003). Modernising the 12-hour shift. *Nursing Standard (Royal College of Nursing (Great Britain): 1987), 17*(19), 33–36. **Reason for exclusion:** Study considered outcomes from a change in shift patterns therefore not related to review objectives.
Rice, S. J. (1992). Occurrence of medication errors on eight and twelve hour shifts. (M.S.N., Bellarmine College). *ProQuest Dissertations and Theses*, **Reason for exclusion:** Study was of insufficient methodological quality e.g. convenience sample, non—validated scale, no information of criteria for error
Richardson, A., Dabner, N., & Curtis, S. (2003). Literature review: Twelve-hour shift on ITU: A nursing evaluation. *Nursing in Critical Care, 8*(3), 103–108. **Reason for exclusion:** Study used number of incidents therefore unable to determine relevance to Benner's taxonomy.
Sullivan, C., & Reading, S. (2002). Nursing shortages: Let's be flexible. *Collegian (Royal College of Nursing, Australia), 9*(4), 24–28. **Reason for exclusion:** Study used number of incidents therefore unable to determine relevance to Benner's taxonomy.
Szczurak, T., Kamińska, B., & Szpak, A. (2007). Estimation of the psychological load in the performance of nurses' work based on subjective fatigue symptoms. *Advances in Medical Sciences, 52 Suppl 1*, 102–104. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Ugrovics, A., & Wright, J. (1990). 12-hour shifts: Does fatigue undermine ICU nursing judgments? *Nursing Management, 21*(1), 64 A. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Vik, A. G., & MacKay, R. C. (1982). How does the 12-hour shift affect patient care? *Journal of Nursing Administration, 12*, 11–14. **Reason for exclusion:** Study did not consider error therefore did not meet review objectives.
Warren, A., & Tart, R. C. (2008). Fatigue and charting errors: The benefit of a reduced call schedule. *Association of Operating Room Nurses. AORN Journal, 88*(1), 88–95. **Reason for exclusion:** Study focus was on implementation of reduced on call hours not shift length therefore did not meet review objectives.

care and paediatric intensive care units (Mills et al., 1983; Ritter, 1981; Scott et al., 2006; Witkoski Stimpfel et al., 2012).

5.5. Interventions

5.5.1. Shift length

Twelve of the studies used shift length as their first intervention with one using work schedule (Trinkoff et al., 2007). Most used 8 h and 12 h as cut off points (Arimura et al., 2010; Mills et al., 1983; Ritter, 1981; Stone et al., 2006; Tanaka et al., 2010; Wilkins and Shields, 2008). Some studies used more specific parameters including <8, 8–9, 9–10, >10 (Dorrian et al., 2006), <8.5, 8.5–12.5, >12.5 (Rogers et al., 2004; Scott et al., 2006), ≥13 h (Trinkoff et al., 2007), 8 h, 12 h, >13 h (Witkoski Stimpfel et al., 2013), 10–11, 12–13, >13 (Witkoski Stimpfel et al., 2012), and ≤8, 8.1–10, 10.1–11.9, 12–13, and >13 (Griffiths et al., 2014).

5.5.2. Error

There were a variety of interventions related to error. These included self-reported error (Arimura et al., 2010; Dorrian et al., 2006; Ritter, 1981; Rogers et al., 2004; Scott et al., 2006; Geiger-Brown and Trinkoff, 2010; Tanaka et al., 2010; Wilkins and Shields, 2008), care left undone (Griffiths et al., 2014), and quality of patient care outcomes (Mills et al., 1983; Stone et al., 2006; Witkoski Stimpfel et al., 2012, 2013). All interventions related to error were able to be categorised according to Benner et al.'s (2002) taxonomy of error and are therefore defined as 'errors'. Table 4 outlines how each of the studies defined error and where this definition fits within Benner's taxonomy.

5.5.3. Associations between 12 h shifts and error

The primary outcome measure considered in this study was the rate of error occurring as a result of working 12 h or more in a single nursing shift in an acute care setting. Four of the studies reported higher rates of error on shifts of up to 8 h (Arimura et al., 2010; Dorrian et al., 2006; Tanaka et al., 2010; Wilkins and Shields, 2008). Arimura et al. and Tanaka et al. both identified higher frequency of self-perceived error-related adverse events among nurses working a three shift system than a two shift system. Dorrian et al. noted shifts of up to 8 h having the highest percentage of error and Wilkins and Shields noted the likelihood of medication error for nurses working 12 h shifts was slightly but significantly lower than for those working shorter shifts.

Table 2a
Quality assessment of comparable cohort/case control studies.

Author	Is the sample representative of patients in the population as a whole?	Are the patients at a similar point in the course of their condition/ illness?	Has bias been minimised in relation to selection of cases and of controls?	Are confounding factors identified and strategies to deal with them stated?	Are outcomes assessed using objective criteria?	Was follow up carried out over a sufficient time period?	Were the outcomes of people who withdrew described and included in the analysis?	Were outcomes measured in a reliable way?	Was appropriate statistical analysis used?
Mills et al. (1983)	No	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes
Trinkoff et al. (2007)	Yes	Yes	N/A	Yes	Yes	N/A	N/A	Yes	Yes

Three studies reported no difference in rates of error between nurses working less than 12 h or more than 12 h shifts (Mills et al., 1983; Ritter, 1981; Stone et al., 2006). Stone et al. identified no statistically significant differences in error rates (risk ratio 1.47, 95% CI 0.90, 2.38).

Six studies reported significant rises in error rates for nurses working greater than 12 h on a single shift. The study by Griffiths et al. (2014) found that nurses working 12 h or more reported higher rates of care left undone than nurses working 8 h or less. Rogers et al. (2004) found the likelihood of making an error was three times higher when nurses worked a shift lasting 12.5 h or more and Scott et al. (2006) found the risk was almost double. Trinkoff et al. (2007) found nurses working greater than or equal to 12 h were significantly more likely to suffer a needlestick injury than those working less than 12 h. In terms of error associated with patient outcomes, Witkoski Stimpfel et al. (2013) reported nurses who worked more than 13 h as significantly more likely than nurses working 8 h to report frequent central line-associated bloodstream infections and Witkoski Stimpfel et al. (2012) reported the higher the proportion of nurses working over 12 h, the less likely pain was controlled.

6. Discussion

This systematic review has examined the relationship between shift length and error amongst nurses working in acute care settings. While earlier studies have examined the relationship between nurse work hours/overtime and nurse and patient outcomes (Bae and Fabry, 2014) and between shift length on quality of patient care (Estabrooks et al., 2009), a specific review of studies examining 12 h shifts, error and nurses has not previously been done. Our findings regarding shift length and error suggest there is a link between the two. This differs from Bae and Fabry who found that while working 12 h shifts negatively affects nurses' health, the evidence linking shift length with patient outcomes was less strong. Our finding also differs from Estabrook et al.'s who were unable to determine a clear finding.

It is important to note that there are multiple ways in which the phenomena of shift length and error can be examined and depending on the lens brought to the literature, findings may differ. In the case of this study, we used a broad definition of error in an attempt to capture the multiple ways in which error can occur in a work setting – such a broad definition is both an advantage (in that it captures multiple variables) and a disadvantage (in that there is an increased likelihood of greater heterogeneity across studies in any meta-analysis). Our meta-analysis did indeed identify high heterogeneity across the five studies we included and we could therefore not draw any definitive conclusions from this analysis of pooled data.

However, supporting our argument that nurses working a shift of 12 h or longer are at increased risk of making an error is the finding that of the 13 studies included in this review, the six that reported significant rises in error rates among nurses working 12 h or more on a single shift (Griffiths et al., 2014; Rogers et al., 2004; Scott et al., 2006; Trinkoff et al., 2007; Witkoski Stimpfel et al., 2012, 2013)

Table 2b
Quality assessment of descriptive/case series studies.

Author	Was study based on a random or pseudo-random sample?	Were the criteria for inclusion in the sample clearly defined?	Were confounding factors identified and strategies to deal with them stated?	Were outcomes assessed using objective criteria?	If comparisons are being made, are there sufficient descriptions of the groups?	Was follow up carried out over a sufficient time period?	Were the outcomes of people who withdrew described and included in the analysis?	Were outcomes measured in a reliable way?	Was appropriate statistical analysis used?
Arimura et al. (2010)	No	No	Yes	Yes	Yes	Yes	N/A	Yes	Yes
Dorrian et al. (2006)	No	Unclear	No	No	No	Yes	N/A	Yes	Yes
Griffiths et al. (2014)	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Ritter (1981)	No	Yes	No	Yes	Yes	No	N/A	Yes	Yes
Rogers et al. (2004)	Yes	Yes	Yes	Unclear	Yes	No	No	Yes	Yes
Scott et al. (2006)	Yes	Yes	N/A	Yes	Unclear	Yes	N/A	Yes	Yes
Stone et al. (2006)	No	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Tanaka et al. (2010)	No	No	Yes	Yes	Yes	Yes	N/A	Yes	Yes
Wilkins and Shields (2008)	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Witkoski Stimpfel et al. (2013)	No	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Witkoski Stimpfel et al. (2012)	No	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes

comprised 89% of the total sample size ($N = 60,780$ with the total sample size $N = 67,967$).

Of the four studies in this review reporting higher rates of error in shifts of up to 8 h, two of these (Arimura et al., 2010; Tanaka et al., 2010) are confounded by the very different work patterns undertaken by the nurses in the study. Confounding factors may also have played a part in the Dorrian et al. (2006) study and the Wilkins and Shields (2008) study. The three studies included in the review that reported no change in rates of error also have limitations. Mills et al. (1983) and Ritter (1981) are both older studies and methodological improvements may impact on the findings of both.

The strong relationship between working 12 h or more on a single shift in an acute care setting and increased rates of error may be associated with variables that have not been examined in this review such as fatigue, levels of stress, handover practices, skill mix or scheduling practices. For example, Dorrian et al. (2006) found sleep duration was a significant predictor of error in their study of shift length and error, and Drach-Zahavy and Hadid (2015) suggest the structure of handover practices is an ideal opportunity to prevent errors. Research examining nurse staffing (skill mix) and patient outcomes demonstrates the higher the education level of nurses, the better the patient outcomes (Aiken et al., 2014) and the higher the number of non-nursing staff providing direct nursing care, the greater the number of medication errors (Seago et al., 2006), also suggesting the existence of mediating variables in the relationship between shift length and error. The relationship between organisational scheduling policies and practices and error in the workplace has not been as well examined although positive nursing practice environments have been positively and significantly associated with medication error interception rates (Flynn et al., 2012). Estryn-Béhar et al. (2012) also strongly recommend work schedules be organised to allow time for shift handover, social support and team building as a means of mitigating the risks of 12 h shifts. We recommend further, specific work be undertaken to examine the strength of the relationships that may exist between variables such as fatigue, stress, skill mix and scheduling practices, shift length and error.

Although other factors may influence the relationship between shift length and rates of error, it is clear from the findings of this review that an association between working 12 h or longer on a single shift in an acute care setting and increased rates of error among nurses remains.

6.1. Implications for practice

This study has significant implications for practice. At present, 12 h shifts are a commonly used approach to rostering systems in many hospital settings. Many nurses themselves are particular advocates of this system, claiming it leads to improved relationships with patients, easier travel to work, better quality time off, and better family time (Estryn-Béhar et al., 2012; O'Connor, 2011; Richardson et al., 2007). However, the findings from this study suggest the risk of error is higher where nurses have worked at least 12 h on a single shift. This suggests that

Table 3
Included studies.

Author country	Method participants setting	Intervention	Outcome measures	Results	Notes
Arimura et al. (2010)	**Method:** Descriptive survey design **Participants:** Nursing staff ($N = 454$) **Setting:** two general hospitals in Japan (338 and 549 beds)	**Intervention A:** Shift system (3 shift system (0830–1700, 1630–0100, 0100–0830) or 2 shift system (1630–0920, including a 2 h nap). **Intervention B:** Self-reported medical error	Occurrence of error	Univariate analysis revealed significant differences between the 'with error' and without error' groups ($p = 0.001$). The odds ratio for nurses working shifts and making an error was 2.1.	Markedly different shift system from that used in Western countries – includes a 2 h regulated sleep during the two-shift system and this may have contributed to lower rates of error among nurses working this system.
Dorrian et al. (2006)	**Method:** Descriptive, exploratory, log book design **Participants:** full time nurses (registered or enrolled) ($N = 23$) **Setting:** Australian metropolitan hospital	**Intervention A:** shift length ($\leq$8, 8–9, 9–10, $\geq$10) **Intervention B:** frequency, type and severity or nursing error or near error	Error/near error occurrence for each shift	Shifts of up to 8 h had the highest percentage of errors. Logistic regression analysis indicated that sleep duration was a significant predictor of error/near error occurrence ($\beta = -0.319$, S.E $> \beta = 0.123$, Wald's $\chi^2 = 6.739$, df = 1, $p = 0.009$).	The majority of 8 h shifts were morning shifts which also had the highest percentage of errors. Variables not entering the equation included shift duration, shift type, number of consecutive shifts, stress ratings. Those having the greatest amount of sleep were least likely to make an error.
Griffiths et al. (2014)	**Method:** Cross sectional survey **Participants:** Registered nurses ($N = 31{,}627$) **Setting:** general medical/surgical units within 488 hospitals across 12 European countries	**Intervention A:** length of shift $\leq$8, 8.1–10, 10.1–11.90, 12–13, >13 **Intervention B:** Care left undone	Frequency of care left undone	Nurses working 12 h or more reported higher rates of care left undone than did nurses working 8 h or less (RR = 1.13, 95% CI = 1.09–1.16). All shifts longer than 8 h were associated with statistically significant increases in the rate of care left undone ($p = < 0.05$).	
Mills et al. (1983)	**Method:** Pilot study – pre and post implementation quality of care data and post implementation survey. **Participants:** professional registered nursing staff ($N = 30$) **Setting:** Surgical Intensive Care Unit	**Intervention A:** implementation of 12 hours shifts **Intervention B:** quality of care data – University of Maryland Hospital, Department of Nursing Process Audit.	Patient care quality scores (University of Maryland Hospital, Department of Nursing Process Audit – scores include documentation of patient care, observation of care related to prevention of injury, protection from infection, special treatments, emotional needs, special procedures)	Analysis using *t* test demonstrated no significant differences between quality of patient care delivered on 8 h or 12.5 h days.	There was a slight decrease in the completeness of admission assessments but that other measures remained constant one year after implementation of the 12 h shift pattern. Improved measures of patient care quality instigated since this study was undertaken may change the result of this study were it to be replicated.
Ritter (1981)	**Method:** Correlational, descriptive study **Participants:** Convenience sample of registered nurses ($N = 83$). **Setting:** ICU and PICU nurses registered in the state of Arizona	**Intervention A:** shift length (combined 12 h and 8 h shift pattern versus traditional 8 h) **Intervention B:** Number of incidents reported on the Incident Report Form	Number of untoward incidents (medication error, procedure error or a patient accident)	No significant difference in the number of patient incidents between either shift pattern (Chi square = 0.19 ($p = 0.34$)	Study undertaken only seven weeks after implementation of new 12 h shift pattern – the short time frame may have contributed to no change in patient care quality scores.

Rogers et al. (2004)	**Method:** Descriptive, exploratory, log book design **Participants:** Registered nurses working full time as hospital staff nurses ($N = 393$) **Setting:** unit-based hospital setting	**Intervention A:** shift length (8.5 and 12.5 h – chosen as '8 h and 12 h shifts are usually scheduled to allow for a half hour hand over period at the end of each shift) **Intervention B:** description of error in logbook	A binary response for making an error during a worked shift	The likelihood of making an error was three times higher when nurses worked shifts lasting 12.5 h or more (odds ration = 3.29, $p = .001$)	
Scott et al. (2006)	**Method:** Descriptive, exploratory, log book design **Participants:** Member of the American Association of Critical Care Nurses, full time employment, registered hospital staff nurse providing direct care to patients ($N = 502$) **Setting:** CCU, ICU or PICU units	**Intervention A:** Actual work shifts, hours (≤ 8.5, >8.5 to <12.5, ≥ 12.5) **Intervention B:** reported error or near error	Risk of error	The risk of making an error almost doubled when nurses worked 12.5 or more consecutive hours (odds ratio 1.94, $p = .03$).	
Stone et al. (2006)	**Method:** Cross sectional survey **Participants:** Direct care hospital nurses ($N = 805$). **Setting:** Thirteen hospitals in New York city.	**Intervention A:** Presence of 12 or 8 h shift pattern. **Intervention B:** Medication events per patient bed, patient falls per patient bed, decubitis ulcer prevalence per patient bed	Unit level quality of patient care	There were no significant differences in any of the quality of patient care variables (Medication events, $p = 0.95$; patient falls $p = 0.64$; decubitis ulcer, $p = 0.66$)	
Tanaka et al. (2010)	**Method:** Descriptive survey design **Participants:** Registered nurses working shifts ($N = 1407$) **Setting:** 5 teaching hospitals in Japan	**Intervention A:** Shift system (two shift – 0800–1700, 1630–0830 with a 2 h nap – and three shift – 0830–1700, 1500–2330, 2300–0900). **Intervention B:** self-perceived error-related adverse events	Frequency of self-perceived error-related adverse events.	There were higher frequencies of adverse events in the three shift system than the two shift system ($p = < 0.05$)	Markedly different shift system from that used in Western countries – includes a 2 h regulated sleep during the two-shift system and this may have contributed to lower rates of error among nurses working this system.
Trinkoff et al. (2007)	**Method:** 3 wave longitudinal survey **Participants:** Actively licensed registered nurses working in nursing for the previous year ($N = 2273$) **Setting:** 2 US States (Illinois and North Carolina)	**Intervention A:** work schedule (hours worked per day ≤ 8, 9–11, ≥ 12) **Intervention B:** Self-reported needlestick injury in the past year	Frequency of needlestick injury	Odds ratio for needlestick injury after working $\geq$ 12 h was 1.68 which was significant at the $p < 0.001$ level (2 tailed test). ≤ 8 h and 9–11 h were insignificant	
Wilkins and Shields, (2008)	**Method:** Cross sectional survey **Participants:** registered nurses providing direct care to hospital patients ($N = 4379$) **Setting:** Canadian hospitals	**Intervention A:** length of shift (8, 12, other) **Intervention B:** medication error	Frequency of wrong medication or dose	For nurses who reported working 12 h shifts the likelihood of medication error was slightly but significantly lower than for those who worked shorter shifts (18% vs 22%). Odds ratio 0.7 (significant to $p < 0.05$)	Nurses working overtime were more likely to make an error

Table 3 (*Continued*)

Author country	Method participants setting	Intervention	Outcome measures	Results	Notes
Witkoski Stimpfel et al. (2013)	**Method:** Secondary analysis of cross sectional survey data **Participants:** Registered nurses providing direct care to paediatric patients ($N = 3710$) **Setting:** 342 acute care hospitals in four states (California, New Jersey, Pennsylvania, Florida)	**Intervention A:** Shift length (8, 12, >13 h) **Intervention B:** central-line-associated bloodstream infections (CLABSI), urinary tract infections	Frequency of central-line-associated bloodstream infections and urinary tract infections	Nurses who worked more than 13 h were significantly more likely than nurses on 8 h shifts to report frequent CLABSI but not UTIs, $p = < 0.05$. CLABSI was reported 2.5 times more frequently by nurses working over 13 h than those working 8	
Witkoski Stimpfel et al. (2012)	**Method:** Secondary analysis of cross sectional survey data **Participants:** Registered nurses providing direct patient care ($N = 22{,}275$) **Setting:** Medical, surgical and intensive care units in 577 hospitals in California, New Jersey, Pennsylvania and Florida	**Intervention A:** length of shift (8–9, 10–11, 12–13, >13 h) **Intervention B:** Pain	Pain was controlled	The higher the proportion of nurses working over 12 h, the less likely pain was controlled (adjusted co-efficient of 0.9 [significant to $p < 0.01$] for > 13 h and 0.2 for nurses working 12–13 h).	

Table 4
Definitions of error study versus Benner et al.'s (2002) taxonomy.

Study	Definition of error	Benner's taxonomy
Arimura et al. (2010)	Committing an incident and/or accident. Includes drug administration errors, incorrect operation of medical equipment, needle stick injuries, surgical errors and patient falls	Medication error Lack of attentiveness Lack of prevention
Dorrian et al. (2006)	'The unintentional use of a wrong plan to achieve an aim, or failure to carry out a planned action as intended' Medical errors – incorrect administration Transcription errors – incorrect transcribing of orders/charts Charting errors Procedural errors Slip or fall Other	Missed or mistaken doctor/ health care provider orders Inappropriate judgement Documentation errors Medication error Lack of prevention
Griffiths et al. (2014)	On your most recent shift, which of the following activities were necessary but left undone because you lacked the time to complete them? Respondents select any items that apply from a list of 13 nursing care activities including: Adequate patient surveillance Documenting nursing care Administering medications Comforting/talking to patients Pain management	Lack of attentiveness Documentation errors Medication error Lack of agency/fiduciary concern Lack of prevention
Mills et al. (1983)	Prevention of injury Protection from infection	Lack of prevention
Ritter (1981)	Any documentable patient care error made by a nurse	All
Rogers et al. (2004)	Any error including: Medication error Procedural error Charting error Transcription error	Medication error Inappropriate judgement Documentation error
Scott et al. (2006)	Medication error Procedural error Charting error Transcription error	Lack of attentiveness Medication error Documentation error Inappropriate judgement
Stone et al. (2006)	Medication events per patient bed Patient falls Decubitis ulcer prevalence per patient bed	Medication error Lack of prevention Lack of intervention on the patient's behalf
Tanaka et al. (2010)	'An unanticipated incident in which the subject nurse made an error which resulted in harm to a patient'	All
Trinkoff et al. (2007)	Needlestick injury	Lack of prevention
Wilkins and Shields (2008)	Medication error	Medication error
Witkoski Stimpfel et al. (2013)	Frequency of central line-associated blood stream infections Frequency of urinary tract infections	Lack of prevention Lack of intervention on the patient's behalf Lack of attentiveness
Witkoski Stimpfel et al. (2012)	Pain was controlled	Lack of intervention on the patient's behalf

hospitals and units currently operating 12 h shift systems should review this approach to scheduling practices due to the potential impact on nurse and patient outcomes. It may be appropriate to consider reducing shift length to 8 or 10 h (although one study (Griffiths et al., 2014) did note that all shifts longer than 8 h were associated with statistically significant increases in the rate of care left undone), or consideration may need to be given to revising scheduling patterns so that significant breaks (for example up to 2 h as in the Arimura et al. (2010) study) be factored into working hours.

6.2. Implications for research

Further research is required to consider factors that may mitigate the risk of error where 12 h shifts are commonly scheduled. For example, longer breaks that include a component of required sleep during night shift may be useful for addressing issues of fatigue that may be a contributing factor to error, however this may not be appropriate during daytime shifts (although longer breaks may be appropriate). Further research is also required to determine what shift patterns are appropriate if 12 h shifts are not. Factors such as fatigue, stress, skill mix and scheduling patterns may also contribute to risk of error and further examination of the influence of these variables on the pathway from shift length to error is also required. We also recommend further research into differing settings. We have focused here on acute hospital settings but suggest that more work is required to examine the links between 12 h shifts and error in mental health, community and aged care settings.

6.3. Limitations of study

As noted earlier, we used a broad definition of error as a means of capturing the multiple ways in which error can occur in a workplace. Benner et al.'s (2002) taxonomy, while

enabling us to be broadly inclusive, may also be considered a weakness of the review as the measurements of error across the studies differ substantially. However, any error will impact on patient safety, quality of care and nurse outcomes and a broad definition can be helpful in determining the overall impact of shift length on these outcomes.

As explained above, variables such as fatigue, stress, skill mix and scheduling practices are not taken into account in this study, and these factors may also have an influence on error rates. Each of the studies also had methodological limitations as noted in Tables 2a and 2b and none were randomised controlled trials.

7. Conclusion

In conclusion, this review shows that rates of error appear higher among nurses working 12 h or more on a single shift in an acute care hospital. Although this review did not explore other potential variables that may mediate or moderate this pathway such as fatigue, stress, skill mix or scheduling practices, it is clear that a relationship does exist. As such, we recommend workplaces review current scheduling practices with a view to limiting shift length to 8–10 h where possible. We also recommend further research into ways in which the risk of error on 12 h or longer shifts can be mitigated.

Acknowledgements

The authors wish to acknowledge the New Zealand Nurses Organisation, the British Commonwealth Nurses War Memorial Fund, and Dr Léonie Walker for her assistance with preparing this manuscript.

Conflict of interest: None.

Funding: New Zealand Nurses Organisation (funded time of first researcher). British Commonwealth Nurses War Memorial Fund (funded JBI systematic review course).

Ethical approval: None.

References

Aiken, L.H., Sloane, D.M., Bruyneel, L., Van den Heede, K., Griffiths, P., Busse, R., Sermeus, W., 2014. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. Lancet 383 (9931), 1824–1830, http://dx.doi.org/10.1016/S0140-6736(13)62631-8.

Arimura, M., Imai, M., Okawa, M., Fujimura, T., Yamada, N., 2010. Sleep, mental health status, and medical errors among hospital nurses in Japan. Ind. Health 48 (6), 811–817.

Bae, S., Fabry, D., 2014. Assessing the relationships between nurse work hours/overtime and nurse and patient outcomes: systematic literature review. Nursing Outlook 62 (2), 138–156.

Benner, P., Sheets, V., Uris, P., Malloch, K., Schwed, K., Jamison, D., 2002. Individual, practice, and system causes of errors in nursing: a taxonomy. JONA 32 (10), 509–523.

Butler, M., Collins, R., Drennan, J., Halligan, P., O'Mathuna, D., Schultz, T., et al., 2010. Hospital nurse staffing models and patient and staff-related outcomes. Cochrane Database Syst. Rev. 7,, http://dx.doi.org/10.1002/14651858.CD007019.pub2 (Art. No. CD007019).

Chen, J., Davis, L., Davis, K., Pan, W., Daraiseh, N., 2011. Physiological and behavioural response patterns at work among hospital nurses. J. Nurs. Manage. 19 (1), 57–68.

Dorrian, J., Lamond, N., van den Heuvel, C., Pincombe, J., Rogers, A.E., Dawson, D., 2006. A pilot study of the safety implications of Australian nurses' sleep and work hours. Chronobiol. Int. 23 (6), 1149–1163.

Drach-Zahavy, A., Hadid, N., 2015. Nursing handovers as resilient points of care: linking handover strategies to treatment errors in the patient care in the following shift. J. Adv. Nurs., http://dx.doi.org/10.1111/jan.12615.

Estabrooks, C., Cummings, G., Olivo, S., Squires, J., Giblin, C., Simpson, N., 2009. Effects of shift length on quality of patient care and health provider outcomes: systematic review. Qual. Saf. Health Care 18, 181–188.

Estryn-Béhar, M., Van der Heijden, B., the Next Study Group, 2012. Effects of extended work shifts on employee fatigue, health, satisfaction, work/family balance, and patient safety. Work 41, 4283–4290.

Flynn, L., Liang, Y., Dickson, G.L., Xie, M., Suh, D., 2012. Nurses' practice environments, error interception practices, and inpatient medication errors. J. Nurs. Scholarsh. 44 (2), 180–186, http://dx.doi.org/10.1111/j.1547-5069.2012.01443.x.

Geiger-Brown, J., Trinkoff, A., 2010. Is it time to pull the plug on 12-hour shifts? Part 1, the evidence. JONA 40 (3), 100–102.

Griffiths, P., Dall'Ora, C., Simon, M., Ball, J., Lindqvist, R., Rafferty, A.M., Schoonhoven, L., Tishelman, C., Aiken, J.L., 2014. Nurses' shift length and overtime working in 12 European countries: the association with perceived quality of care and patient safety. Med. Care 52 (11), 975–981, http://dx.doi.org/10.1097/MLR.0000000000000233.

Hodgkinson, B., Koch, S., Nay, R., Nichols, K., 2006. Strategies to reduce medication errors with reference to older adults. JBI Libr. Syst. Rev. 4 (1), 2–41.

Joanna Briggs Institute, 2011. Reviewers Manual. Joanna Briggs Institute, Adelaide, Australia.

Johnstone, M., Kanitsaki, O., 2006. The ethics and practical importance of defining, distinguishing and disclosing nursing errors: a discussion paper. Int. J. Nurs. Stud. 43, 367–376.

Joyce, K., Pabayo, R., Critchley, J., Bambra, C., 2010. Flexible working conditions and their effects on employee health and wellbeing. Cochrane Database Syst. Rev. (2.) , http://dx.doi.org/10.1002/14651858.CD008009.pub2 (Art. No.: CD008009).

Ker, K., Edwards, P., Felix, L., Blackhall, K., Roberts, I., 2010. Caffeine for the prevention of injuries and errors in shift workers. Cochrane Database Syst. Rev. 5,, http://dx.doi.org/10.1002/14651858.CD008508 (Art. No. CD008508).

Mills, M.E., Arnold, B., Wood, C.M., 1983. Core-12: a controlled study of the impact of 12-hour scheduling. on personnel and patient care. Nurs. Res. 32 (6), 356–361.

O'Connor, T., 2011. Director of nursing at ADHB wants end to 12-hour shifts. Kai Tiaki Nursing New Zealand 17 (9), 37.

Richardson, A., Turnock, C., Harris, L., Finley, A., Carson, S., 2007. A study examining the impact of 12-hour shifts on critical care staff. J. Nurs. Manage. 15 (8), 838–846.

Ritter, L., 1981. Effects of the combined 12–8 and 8-hour shifts on nursing and patient outcomes. The University of Arizona, United States – Arizona.

Rogers, A.E., Wei-Ting, H., Scott, L.D., Aiken, L.H., Dinges, D.F., 2004. The working hours of hospital staff nurses and patient safety. Health Aff 23 (4), 202–212.

Seago, J., Williamson, A., Atwood, C., 2006. Longitudinal analyses of nurse staffing and patient outcomes: more about failure to rescue. JONA 36 (1), 13–21.

Scott, L.D., Rogers, A.E., Hwang, W., Zhang, Y., 2006. Effects of critical care nurses' work hours on vigilance and patients' safety. Am J Crit Care 15 (1), 30–37.

Stone, P.W., Du, Y., Cowell, R., Amsterdam, N., Helfrich, T.A., Linn, R.W., et al., 2006. Comparison of nurse, system and quality patient care outcomes in 8-hour and 12-hour shifts. Med. Care 44 (12), 1099–1106.

Tanaka, K., Takahashi, M., Hiro, H., Kakinuma, M., Tanaka, M., Kamata, N., et al., 2010. Differences in medical error risk among nurses working two- and three-shift systems at teaching hospitals: a six-month prospective study. Ind. Health 48 (3), 357–364.

Trinkoff, A.M., Le, R., Geiger-Brown, J., Lipscomb, J., 2007. Work schedule, needle use, and needlestick injuries among registered nurses. Infect. Control Hosp. Epidemiol. 28 (2), 156–164.

Wagstaff, A., Sigstad, L., 2011. Shift and night work and long working hours – a systematic review of safety implications. Scand. J. Work Environ. Health 37 (3), 13–185.

Wilkins, K., Shields, M., 2008. Correlates of medication error in hospitals. Health Reports/Statistics Canada, Canadian Centre for Health Information.(Rapports sur la santé/Statistique Canada). Centre canadien d'information sur la santé 2008 June 19 (2), 7–18.

Witkoski Stimpfel, A., Lake, E., Barton, S., 2013. How differing shift lengths relate to quality outcomes in paediatrics. JONA 43 (2), 95–100.

Witkoski Stimpfel, A., Sloane, D., Aiken, L., 2012. The longer the shifts for hospital nurses, the higher the levels of burnout and patient dissatisfaction. Health Aff. 31 (11), 2501–2509.