Outcomes sensitive to critical care nurse staffing levels: A systematic review

Pamela J. L. Rae*,1, Susie Pearce1, P. Jane Greaves2, Chiara Dall’Ora3, Peter Griffiths3
Ruth Endacott1,4,5

1. School of Nursing & Midwifery, University of Plymouth, Plymouth PL4 8AA
2. School of Health and Life Sciences, University of Northumbria, Newcastle Upon Tyne UK
3. School of Health Sciences, University of Southampton
4. Royal Devon and Exeter Hospital/University of Plymouth Clinical School, Royal Devon and Exeter Hospital, Barrack Road Exeter EX2 5DW
5. School of Nursing & Midwifery, Monash University, Melbourne, Vic 3199

*Corresponding author:
pamela.rae@plymouth.ac.uk, ORCID 0000_0003_0370_6570
Dr Pamela J L Rae
South West Clinical Schools
School of Nursing and Midwifery
Faculty of Health
University of Plymouth
Rm 104, 8 Kirkby Place
Drake Circus
Plymouth
Devon
PL4 8AA
01752 583212

Declaration of Interest: Funding Source and Role of Funding Source

This paper presents independent research funded by the National Institute for Health Research (Programme Development Grants, Safe staffing in ICU: development and testing of a staffing model, NIHR200100). The views expressed in this publication are those of the author(s) and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care, neither of whom have had involvement in any aspect of the design, data collection, synthesis, interpretation or writing of, this review.

Acknowledgement

We would like to thank members of the SEISMIC (Study to Evaluate the Introduction of nurse Staffing Models in Intensive Care) study team for their feedback during the review process: Dr Suzanne Bench, Mrs Carole Boulanger, Miss Fiona Boyd, Mrs Jennifer Gordon, Dr Doug Gould, Prof David Harrison, Dr Jeremy Jones, Dr Thomas Monks, Mr Paul Mouncey, Prof Natalie Pattison, Dr Annette Richardson, Prof Kathryn Rowan.
Abstract

Objective To determine associations between variations in registered nurse staffing levels in adult critical care units and outcomes such as patient, nurse, organisational and family outcomes.

Methods: We published and adhered to a protocol, stored in an open access repository and searched for quantitative studies written in the English language and held in CINAHL Plus, MEDLINE, PsycINFO, SCOPUS and NDLTD databases up to July 2020.

Three authors independently extracted data and critically appraised papers meeting the inclusion criteria. Results are summarised in tables and discussed in terms of strength of internal validity. A detailed review of the two most commonly measured outcomes, patient mortality and nosocomial infection, is also presented.

Results: Our search returned 7,960 titles after duplicates were removed; 55 studies met the inclusion criteria. Studies with strong internal validity report significant associations between lower levels of critical care nurse staffing and increased odds of both patient mortality (1.24 to 3.50 times greater) and nosocomial infection (3.28 to 3.60 times greater), increased hospital costs, lower nurse-perceived quality of care and lower family satisfaction. Meta-analysis was not feasible because of the wide variation in how both staffing and outcomes were measured.

Conclusions A large number of studies including several with high internal validity provide evidence that higher levels of critical care nurse staffing are beneficial to patients, staff and health services. However, inconsistent approaches to measurement and aggregation of
staffing levels reported makes it hard to translate findings into recommendation for safe staffing in critical care.

**MeSH Keywords:** Critical care; cross infection; health care; health workforce; mortality; outcome assessment; registered nurse.

**Implications for Clinical Practice**

- Reduced critical care nurse staffing is associated with increased rates of patient mortality and increased risks of nosocomial infection.

- Reduced staffing may also be associated with increased hospital costs, lower nurse-perceived quality of care and lower family satisfaction.

- Inconsistent approaches to measurement and aggregation of staffing levels reported in research makes it hard to translate findings into recommendation for safe staffing in critical care.
Introduction

Studies show that nurse staffing levels in acute hospital wards have serious implications for patient outcomes and staff wellbeing (Ausserhofer et al., 2014; Griffiths, Ball, Murrells, Jones, & Rafferty, 2016). For example, nurse staffing has been linked to patient mortality, failure to rescue, and staff satisfaction (Aiken et al., 2012; Aiken et al., 2014). Getting the level of nurse staffing right was brought into sharp focus by the UK Francis report into failings at an NHS hospital (Francis, 2013) and acknowledgement of the need to improve the safety of patients and quality of care (Berwick, 2013).

Systematic reviews of research in this field show clear links between nurse staffing and patient outcomes (Assaye, Wiechula, Schultz, & Feo, 2020; Griffiths et al., 2014). However, these reviews are largely concerned with staffing levels in general acute wards and not in intensive care settings. Nurse staffing levels in intensive care settings are modelled on different indices to those in acute settings and are generally much higher, reflecting that patients with higher acuity need more nursing resource. For example, the UK classifies higher acuity patients (Level 3 patients) as requiring a minimum 1:1 registered nurse:patient ratio, whereas high dependency patients recovering from critical illness (Level 2 patients) require a minimum of 1:2 to deliver direct care (Bray et al., 2010). Thus in this example, unlike acute care settings, staffing levels in UK intensive care tend to follow fixed ratios and are potentially impacted in a different way when nurse workloads increase due to changes in treatment policy or the clinical profile of patients admitted. Therefore, the link between variations in nurse staffing level and outcome, evident in the acute care setting, may not replicate in critical care settings.

Existing reviews on nurse staffing levels have excluded studies carried out exclusively in settings used for the delivery of critical care (Griffiths et al., 2014) or have focussed on other
care settings such as acute or paediatric care (Kim et al., 2018; Wilson, Bremner, Hauck, & Finn, 2011). Previous reviews examining critical care nurse staffing have focussed on patient outcomes (Liang et al., 2010; Numata et al., 2006; Penoyer, 2010; West, Mays, Rafferty, Rowan, & Sanderson, 2009) were carried out a decade or more ago and research in this area has since grown. Therefore, this review will synthesise data from studies examining associations between outcomes and variations in critical care nurse staffing levels in Intensive Care Units (ICUs).

**Review question**

The objective of this review is to identify associations between variations in registered nurse staffing in adult critical care and patient, nurse, organisational and family outcomes.

**Methods**

This is a systematic review. We published and adhered to a protocol, stored in an open access repository, Plymouth Electronic Archive and Research Library (PEARL) (https://pearl.plymouth.ac.uk/handle/10026.1/16017).

**Search strategy**

We searched CINAHL Plus, MEDLINE, PsycINFO and SCOPUS for published studies and the Networked Digital Library of Theses and Dissertations (NDLTD) and conference proceedings listed in SCOPUS for unpublished studies and grey literature. We used two categories of search terms, one to identify staffing, for example, ‘nurse staffing’ and the other to identify specialism, for example, ‘intensive care’. An example of our MEDLINE search strategy is included in supplementary material (e-component file 1).
Study selection

We included primary quantitative and mixed methods studies exploring the association between variations in staffing levels of registered nurses working in adult critical care units, and outcomes. We use the term Registered Nurse (RN) to include all nursing staff who were professionally registered or licensed (for example, Licensed Professional Nurses).

We included studies measuring any outcome, for example, patient, family, staff, care quality or organisational outcomes. The term ‘family member’ is defined as per Rowan et al (2014) in their study evaluating family satisfaction with adult critical care: “a person with a close familial, social or emotional relationship with the patient and is not restricted solely by next of kin” (pg 12) (Wright et al., 2015).

We did not include studies conducted in paediatric or neonatal intensive care units (PICUs, NICUs) because nursing in these environments differs in terms of, for example, parental involvement and care delivery (Aiken et al., 2014; Nipshagen, Polderman, DeVictor, & Gemke, 2002). Studies conducted in a mixture of NICUs, PICUs and adult ICUs were only included if their analysis for adult ICU was reported separately.

Studies published in English, with no lower date limit up to July 2020, were included. In line with guidance from the Cochrane Handbook (Lefebvre et al., 2019), we imposed no a-priori date limits. Citations were managed using Endnote X8 (Clarivate Analytics, PA, USA) and uploaded to RAYYAN systematic review software (Qatar Computing Research Institute). The results of the search are reported at Figure 1 (Tricco et al., 2018).
Data extraction and quality assessment

Data was extracted from quantitative and mixed methods (quantitative component only) studies using a template developed by NICE (Griffiths et al., 2014) and presented in e-component file 1. Although we intended to carry out a meta–analysis of data on effect sizes, this was not possible because there was considerable heterogeneity in how staffing levels and outcomes were measured in included studies. As an example of the former, Table 1 presents a selection of nurse staffing level operational definitions. As an example of the latter (see e-component file 2), the range of operational definitions of patient mortality included in-hospital mortality, ICU mortality, 28- and 30-day mortality. As a quality assessment, the internal and external validity of each study was evaluated independently by two reviewers against a framework adapted from NICE for observational/cross-sectional studies (Griffiths et al., 2014). The framework (e-component file1) assigns each study one of three grades based on methodological rigour: strong, moderate or weak. Whether included studies explored nurse-staffing and outcomes as primary or secondary aims, assessment of methodical rigour remained focussed on how these data were collected and analysed. The main focus for the review was on internal validity however, we also present our assessment of external validity. Internal validity was assessed primarily by rating the design of the study on aspects such as when and how data were collected, the reliability of measures used and level of risk adjustment incorporated into analysis. Strong internal validity therefore suggests that the statistical conclusions of a study are likely to be an unbiased reflection of the association between staffing variation and outcome in the setting and population studied. The quality appraisal checklist is provided as e-component file 1.
Data synthesis and integration

Data extraction was conducted for all studies, regardless of their methodological quality however, our synthesis is focussed on the findings of strong internal validity studies. Therefore, while moderate and weak studies are included in the review with regard to the description of, for example, number of studies investigating each outcome, their findings are presented in tables only.

We gave particular emphasis to studies where staffing and outcomes were measured simultaneously or longitudinally as opposed to cross-sectional studies where average outcomes were associated with average staffing over long intervals. Although the former type of studies do not provide evidence of cause, they do allow an exploration of the temporal link between staffing variation and outcome. We refer to these studies as longitudinal.

Results

The electronic search yielded 10,882 citations; duplicate removal left 7,960 titles of which 55 met the inclusion criteria (Figure 1). The earliest date of publication was 1993. Study designs comprised 41 cross-sectional (including retrospective and prospective aspects) and 14 longitudinal studies (8 retrospective and 6 prospective).

The number of centres per study ranged from a single ICU to 1,265 ICUs. The number of patients per study varied between 30 and 159,400. Forty-one studies measured patient outcomes (see Table 2) which were predominantly mortality and nosocomial infection (see e-component file 2 for detailed evidence summaries). Thirteen studies measured care process outcomes, including costs (Table 3). Four studies measured family satisfaction (Table 4), and
one study measured nurse outcomes (Table 4). Forty-four studies were designed with nurse staffing as a primary focus, twenty-three studies (42%) had strong internal validity (++). We classified 14 studies as longitudinal (25%), of which six had strong internal validity (43%) and eight had moderate internal validity (57%). Studies with high internal validity all had some form of risk-adjustment to account for potential confounders.

**Measures of nurse staffing**

Although the precise definition of ‘nurse’ was not specified in all studies, studies with nurses and nursing assistants included but not delineated in nurse staffing calculations, were excluded. Studies conducted in countries where critical care nurse:patient ratio has been previously defined in professional guidance, and such definition clearly refers to Registered Nurse, were assumed to have used the term in the same way.

Staffing levels were measured and defined in a multitude of ways. Some form of Nurse:Patient ratio (N:P) was used in 39 (71%) studies. Other methods included: Bed to Nurse ratio (Checkley et al., 2014; Faisy et al., 2016; Kim¹, Kim, & Shin, 2019; Kim², Kim, & Lee, 2020; Seynaeve et al., 2011; West et al., 2014); Nursing Hours Per Patient Day (Blegen, Goode, Spetz, Vaughn, & Park, 2011; Boev & Xia, 2015; Fridkin et al., 1996; Stone et al., 2007; Van den Heede et al., 2009); number of nurses (Dancer et al., 2006). Six studies incorporated a measure of workload using, for example, the Nursing Activities Score (Gerasimou-Angelidi, Myrianthefs, Chovas, Baltopoulos, & Komnos, 2014; Strazzieri-Pulido, González, Nogueira, Padilha, & Santos, 2019) or the Therapeutic Intervention Scoring System (Lee et al., 2017). One study measured nurses’ perceptions of staffing adequacy alongside N:P ratios (Cho et al., 2009). There was variation in the method of collection and calculation of staffing level data. For example, some N:P ratios were extracted from hospital databases, others were collected through surveys of ICU ward managers. The
calculation of N:P sometimes reflected a day shift but in other instances, reflected an average of three shifts/24 hours. Staffing might be treated as a fixed, static characteristic of the unit and compared to other units with a different fixed ratio. Alternatively, staffing might be calculated during the unique study period and recorded per shift alongside the outcome measure. All studies measured natural variation in nurse staffing levels rather than an intended change in nurse staffing levels designed to improve outcomes.

**INSERT TABLE 1 HERE**

**Patient Outcomes**

Table 2 provides a summary of the 41 patient outcome studies (strong, moderate and weak) meeting eligibility criteria for the review.

**INSERT TABLE 2 HERE**

The patient outcomes measured across the 41 studies were mortality (n=21 studies), nosocomial infection (n= 13), adverse events (n= 8), length of hospital stay (n= 5), length of ICU stay (n= 3), number of ventilator days (n= 2), length of weaning (n= 1), multiple organ failure (n= 1) and patient satisfaction (n= 1). The studies measuring mortality and nosocomial infection are presented in detail in supplemental information tables (e-component file 2). Of the 21 studies measuring mortality, 17 measured hospital mortality, four measured 28- or 30-day mortality and three measured ICU mortality.

Six studies of strong internal validity used multiple categorises to describe staffing levels (for example, highest level of staffing, second highest, lowest), however, categories were defined with different thresholds (Kim² et al., 2020; Margadant et al., 2020; Neuraz et al., 2015; Sakr et al., 2015; Stone et al., 2007; Tarnow-Mordi, Hau, Warden, & Shearer, 2000). Two of the
studies (Kim² et al., 2020; Neuraz et al., 2015) found a non-significant increase in magnitude of staffing benefit with each increase in staffing level (please see e-component file 2).

There were fifteen studies with high internal validity measuring mortality outcomes. Ten of the fifteen, including three with a longitudinal design (Lee et al., 2017; Neuraz et al., 2015; Tarnow-Mordi et al., 2000), report evidence of a statistically significant association between higher staffing levels and lower mortality rates. However, one study with high internal validity showed a significant association between higher staffing and higher mortality (Dodek et al., 2015).

Five studies rated with strong internal validity explored associations between nurse staffing and nosocomial infection (Amaravadi et al., 2000; Blegen et al., 2011; Halwani et al., 2006; Hugonnet et al., 2007; Stone et al., 2007). Two of these studies incorporated a longitudinal design (Halwani et al., 2006; Hugonnet et al., 2007) and all five report statistically significant associations between higher levels of staffing and lower rates of infection.

Four of eight patient outcomes studies investigating adverse events had strong internal validity, all four report statistically significant associations between increased staffing and reduced rates of adverse events (Amaravadi et al., 2000; Dang et al., 2002; Dimick et al., 2001; Stone et al., 2007).

Three studies with strong internal validity measured hospital length of stay (Amaravadi et al., 2000; Dimick et al., 2001; Dodek et al., 2015). One study reported a statistically significant association between higher staffing and shorter length of stay (Amaravadi et al., 2000) and one reported a non-significant beneficial association (Dimick et al., 2001). However, one reported a significant detrimental association, with higher staffing associated with longer stays (Dodek et al., 2015).
Two of three studies measuring ICU length of stay were rated with strong internal validity (Dodek et al., 2015; Verburg et al., 2018). While the latter found a significant association between higher staffing and reduced length of stay, the former found an inconsistent relationship with length of stay both in magnitude and direction across regions in the study.

Other patient outcome studies with strong internal validity found a significant association between higher staffing levels, and reduced length of weaning (Thorens et al., 1995); reduced number of ventilator days (Stone et al., 2007) and reduction in multiple organ failure (Jansson et al., 2020).

**Care process outcomes and costs**

Table 3 provides a summary of all care process outcomes and costs studies (strong, moderate and weak) meeting the eligibility criteria for inclusion in the review.

*INSERT TABLE 3 HERE*

Four studies were rated to have strong validity, two of which were large, multi-centre studies measuring hospital costs (Amaravadi et al., 2000; Dimick et al., 2001). Both found significant associations between higher nurse staffing and reduced patient hospital cost. Neither of the studies incorporated temporal measures.

The third study used both nurse to patient ratio and nurses’ perception of adequate staffing, as measures of staffing levels (Cho et al., 2009). Both staffing measures showed a statistically significant association between higher staffing and higher nurse-perceived quality of care.

The fourth study reported significant associations between higher staffing levels and timely decisions on life saving treatment (Azoulay et al., 2009).
Nurse outcomes and family satisfaction

Table 4 provides a summary of nurse and family outcome studies (strong, moderate and weak) meeting the eligibility criteria for inclusion in the review.

Two studies were rated to have strong internal validity. One study explored the association between nurse staffing and nurse outcome (Cho et al., 2009) and found statistically significant beneficial associations between subjective measures of staffing and burnout, job-satisfaction and intention to leave. An objective measure of staffing also showed beneficial (but not statistically significant) associations with job-satisfaction and intention to leave, but showed a detrimental (non-significant) association with burnout. The second study (Stricker et al., 2009) showed a significant association between higher staffing levels and greater family satisfaction.

Discussion

We reviewed 55 papers and found evidence from studies assessed to have strong internal validity, that higher levels of critical care nurse staffing are beneficial to a range of outcomes. In particular, there is strong evidence that higher critical care nurse staffing levels are associated with reduced rates of mortality and nosocomial infection. There is also strong evidence of associations with reduced hospital costs, increased family satisfaction, reduced nurse burnout, increased nurse job satisfaction and less intention to leave the nursing profession.

Studies assessed to have moderate or weak internal validity generally supported these findings (please refer to Tables 2 to 4 for details). These studies also showed that higher
staffing tended to be associated with increased adherence to clinical protocols, reduced length of ventilator days and weaning, and impacted care processes such as facilitating time to make decisions on life-forgoing treatment or using physical restraints.

Evidence of an association between higher staffing and reduced length of hospital or ICU stay however, was mixed. Regardless of risk adjustment, there may be organisational factors, such as diversity in the organisation of intermediate care, and hospital capacity for recovering ICU patients, that may affect outcomes such as hospital length of stay. Whilst longer ICU stay is acknowledged as a poor outcome, hospital length of stay may mean better/prolonged access to rehabilitation (therapists and personnel). Hence, wide variation in the impact of increased staffing on hospital length of stay (from significantly detrimental to significantly beneficial) should be interpreted with caution.

Thirty three of the 55 studies included in our review were published since the last systematic reviews on patient outcomes in this field (Numata et al., 2006; West et al., 2009), indicating the enduring interest in this as a research focus. The patient outcome themes in our review are broadly similar to those in previous reviews however, our review presents a stronger body of evidence (more studies and stronger internal validity) and incorporates a broader range of outcomes.

Differences in ICU staffing levels across countries, and particularly between countries, persist across reviews and are reflected in other multi-country studies. For example, researchers examining early mobilisation practices contrasted levels of staffing across both the UK and France; differences in levels were mostly inter (rather than intra-) country with 97% of UK ICUs reporting a 1:1 nurse:patient ratio, and 90% of French ICUs reporting ratios of 1:3 or higher (Bakhru et al., 2016). Aspects of ICU nursing practice such as patient and
family engagement (for example use of ICU diaries and open visiting for families) have been reported as consistent within, but variable between, countries (Kleinpell et al., 2018). A survey of 24 European countries revealed large variation in availability, duration and qualification awarded for critical care nursing programmes (Endacott et al., 2015). These studies indicate that ICU nurses are perhaps undertaking different roles in different countries so any attempt to identify an optimal N:P ratio for international application would be doomed to failure.

**Limitations**

As with all reviews, there are limitations to our conclusions. The methodological quality of included studies, for example, determines the confidence with which we offer our final summary. We assessed two fifths of studies to have strong methodical rigour with respect to measuring and analysing nurse staffing and outcome variables. We used a pre-established detailed framework replete with numerous examples of poor, moderate and strong grades to guide us in our assessment (please see e-component file 1). In addition, our decision to assign separate grades for internal and external validity reduces the number of conflating issues. Although this system worked well in so far as there was very little discord among assessors’ grade assigning, we acknowledge that there remains an unavoidable degree of subjectivity in establishing a cut-off for each of the three grades. Nevertheless, at the very least, this grading system enables a relative distinction of rigour.

An additional limitation centres on the heterogeneity of risk adjustment methods used across included studies. Studies assessed with strong internal validity all reported key patient/nurse and unit/hospital risk adjustments. This included measuring and accounting for a plethora of variables such as nurse specialist training, length of critical care nursing experience, patient age, sex, type of admission, ICU specialism, type of hospital (public, private, large, small) and many more, too numerous to list here. Therefore, while we are
confident of the methodological rigour of studies we assessed as strong, we have not attempted to compare data based on the specific type of risk adjustments carried out, some of which were idiosyncratic to a particular study. In summary, strong studies all carried out risk adjustments but not necessarily with the same combination of confounders. However, we would undermine our endeavour to synthesise these data should we attempt to group studies by both outcome measured and by type and combination of risk adjustment used.

We acknowledge that in terms of time periods, older compared to newer studies may represent data collected against a different backdrop of intensive care environment and staffing. However, this difference in backdrop also applies to studies conducted within the same time period: differences between two studies conducted in 2019 may be just as likely as a 2000 study and a 2019 study. This is largely because of variations in the use (rather than definition) of high intensity care wards. The stated function of the wards as, for example, ICU, HDU, Cardiac, Burns, Neuro, represents variation in use and throughput that is inconsistently available in studies. Further, the number of ICU beds per 100 hospital beds varies greatly. The more abundant the supply of ICU beds the more low-acuity patients likely to be admitted, both by beds being available for more precautionary admissions and by patients being held in ICU longer following serious illness. One way to address some of these issues would be to directly compare only studies using the same risk adjustments. However, as already discussed, this was not feasible. Therefore, we saw no justification in synthesising data by time period, and in line with systematic review guidance (Lefebvre et al., 2019), we did not impose a lower date limit on included studies.

Overall, the heterogeneity of approaches to measurement in these studies suggests there needs to be a more coherent framework with which to examine the impact of nurse staffing. As a case in point, we were unable to identify staffing level thresholds because there was no common framework and as such, the optimal ‘dose’ of nursing remains unknown. A
previous review of nurse staffing (Greaves et al., 2018) reported that formulae used to calculate staffing were used inconsistently and that the clinical judgment of a senior nurse performed as well as the best scoring system. This was reflected in the large multi-centre study in which N:P ratios and perceptions of adequate staffing were both significantly associated with perceived quality of care (Cho et al., 2009).

**Conclusion.**

A large number of studies including several with large samples and with strong internal validity provide clear evidence that higher levels of ICU registered nurse staffing are beneficial, particularly to patient survival. There is a more mixed picture, with fewer studies or weaker evidence, of protection against iatrogenic events, family satisfaction, nurse well-being and hospital costs. Heterogeneity across research methodologies, with inconsistent approaches to measurement and aggregation of staffing levels, makes it hard to translate findings from this body of evidence into recommendations for safe staffing in critical care. However, treating staffing as a study variable rather than a demographic characteristic of the ICU, would focus attention on how it is measured and enable researchers to better develop or evaluate changing models of staffing. Evidence of benefit is also tempered by the absence of intervention studies in which the dose of nurse staffing was adjusted in order to measure impact on outcomes.
Figure 1 and Tables 1-4

Figure 1: Flowchart of study inclusion

Records identified through database search  
$n = 10,882$

2,912 duplicates removed

Records screened in title and abstract  
$n = 7,960$

7,831 records excluded

Full-text articles assessed for eligibility  
$n = 129$

74 records excluded

35 No adequate measure or no measure of staffing level variation  
14 ICU data not analysed separately from other units’ data  
10 RN data not analysed separately from other staff’s data  
6 Data presented in another, included, study  
3 No response from author over methodological query  
2 RN data were not from ICU RNs  
1 Staffing data were not from RNs  
1 Missing analysis  
1 No measure of outcome  
1 Commentary

Studies included in the review  
$n = 55$
<table>
<thead>
<tr>
<th>Operational Definition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:P per eight-hour shift (Verburg et al., 2018; Vicca, 1999)</td>
<td></td>
</tr>
<tr>
<td>Total number of nurses working/patient census for that day (Hugonnet et al., 2007)</td>
<td></td>
</tr>
<tr>
<td>Collective N:P of ratios taken at two separate one hour time points on one day (Sakr et al., 2015)</td>
<td></td>
</tr>
<tr>
<td>Self-reported number of patients each nurse had cared for on their last shift, from which a mean critical care staffing measure for each participating hospital was generated (Kelly et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Fixed ICU characteristic with no evidence that the intended staffing (e.g. 1:1 N:P) was met at the time of outcome measurement (Dodek et al., 2015; Johannessen et al., 2011; Pronovost et al., 1999)</td>
<td></td>
</tr>
<tr>
<td>First Author (Year) Country</td>
<td>Design</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Baykara (2018) Turkey</td>
<td>CS</td>
</tr>
<tr>
<td>Blegen (2011) US</td>
<td>CS</td>
</tr>
<tr>
<td>Blot (2011) Europe</td>
<td>PO</td>
</tr>
<tr>
<td>Boev (2015) US</td>
<td>CS, RO</td>
</tr>
<tr>
<td>Checkley (2014) US</td>
<td>CS</td>
</tr>
<tr>
<td>Chittawatananarat (2014) Thailand</td>
<td>RO</td>
</tr>
<tr>
<td>Cho (2008) South Korea</td>
<td>RO</td>
</tr>
<tr>
<td>Dancer (2006) UK</td>
<td>RO, L</td>
</tr>
<tr>
<td>Dimick (2001) US</td>
<td>RO</td>
</tr>
<tr>
<td>Dodek (2015) Canada</td>
<td>CS, RO</td>
</tr>
<tr>
<td>Dorsay (2000) US</td>
<td>RO, L</td>
</tr>
<tr>
<td>Faisley (2016) France</td>
<td>PO, L</td>
</tr>
<tr>
<td>Fridkin (1996) US</td>
<td>RO, L</td>
</tr>
<tr>
<td>Graf (2010) Germany</td>
<td>CS</td>
</tr>
<tr>
<td>Halwani (2006) UK</td>
<td>PO, L</td>
</tr>
<tr>
<td>Hugonnet (2007) Switzerland</td>
<td>PO, L</td>
</tr>
<tr>
<td>Jansson (2020) Finland</td>
<td>CS</td>
</tr>
<tr>
<td>Jansson (2019) Finland</td>
<td>PO, L</td>
</tr>
<tr>
<td>Johannessen (2011) Norway</td>
<td>CS</td>
</tr>
<tr>
<td>Kelly (2014) US</td>
<td>CS</td>
</tr>
<tr>
<td>First Author (Year) Country</td>
<td>Design</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Kim (2020) South Korea</td>
<td>CS</td>
</tr>
<tr>
<td>Kim (2019) South Korea</td>
<td>CS, RO</td>
</tr>
<tr>
<td>Kim (2012) South Korea</td>
<td>PO</td>
</tr>
<tr>
<td>Lee (2017) Hong Kong</td>
<td>RO, L</td>
</tr>
<tr>
<td>Margadant (2020) Holland</td>
<td>CS</td>
</tr>
<tr>
<td>Pronovost (1999) US</td>
<td>RO, PO</td>
</tr>
<tr>
<td>Schwab (2012) Germany</td>
<td>RO, L</td>
</tr>
<tr>
<td>Seynaeve (2011) Holland</td>
<td>CS, RO</td>
</tr>
<tr>
<td>Stone (2007) US</td>
<td>CS, PO</td>
</tr>
<tr>
<td>Strazzieri-Pulido (2019)</td>
<td>RO</td>
</tr>
<tr>
<td>Tarnow-Mordi (2000) UK</td>
<td>RO, L</td>
</tr>
<tr>
<td>Thorens (1995) Switzerland</td>
<td>PO, L</td>
</tr>
<tr>
<td>Vicca (1999) UK</td>
<td>RO, L</td>
</tr>
<tr>
<td>West (2014) UK</td>
<td>CS, RO</td>
</tr>
<tr>
<td>Yeh (2004) Taiwan</td>
<td>CS</td>
</tr>
</tbody>
</table>

**Design**: CS cross-sectional, RO retrospective observational, PO prospective observational, L longitudinal. **Nurse staffing measure** [Ratio antecedents and consequents may have been calculated either way round, for e.g. as N:P or P:N: N:P Nurse to Patient ratio (Table 1 breaks this down further); N:P in tertiary hospitals, N:P in secondary hospitals (Cho, 2008); NNPR Night-time Nurse to Patient Ratio; N: VentP nurse to ventilated patient ratio; N: Bed Nurse to bed ratio; NAS Nurses Activity Score; N number of nurses on shift; NHPPD nursing hours per patient day; NNR NAS score per nurse ratio; Workload composite measure based on average nursing requirement per occupied bed and peak occupancy in any shift during patient’s stay (Tarnow-Mordi, 2000); total TISS-76 divided by the average number of direct patient care nurses per 24hour day (Lee, 2017). **Staffing quotient**: number of nurses (1:1 patients = 1) + (1:2 patients = 0.5), calculated twice a day at the start of each 12 hour shift (Dorsey, 2000). NR – Not Reported. **Number of ICUs**: 2 37 ICUs in British Columbia; 20 ICUs in Ontario. **Internal, External Validity**: ++ Strong; + Moderate; - Weak. **Adverse events**: Risk of reintubation (Amaravadi, 2000; Dimick 2001); pressure injury (Stone, 2007; Strazzieri-Pulido 2019); complications during surgery (Dang, 2002); unexpected cardiac arrest, unplanned extubation (Fairy, 2016; Yeh, 2004), reintubation after planned extubation, readmission with 48 hrs (Faisy, 2016); adverse drug events (Seynaeve, 2011). **LoS Length of stay**: ▲ Higher staffing is significantly (p<0.05) beneficial ▲ Higher staffing is numerically beneficial. ▼ Higher staffing is significantly (p<0.05) detrimental. ▼ Higher staffing is numerically detrimental. ☐ No evidence of a significant association (p≥0.05) and no figures given.
Table 3: Association between nurse staffing level and care process outcomes, and costs

<table>
<thead>
<tr>
<th>First Author (Year) Country</th>
<th>Design</th>
<th>Nurse Staffing Measure</th>
<th>Number of RNs</th>
<th>Number of Patients</th>
<th>Number of ICUs</th>
<th>Internal Validity</th>
<th>External Validity</th>
<th>Adequate Pain Management</th>
<th>Professional Collaboration</th>
<th>DFST</th>
<th>Hospital costs</th>
<th>Nurse perceived quality of care</th>
<th>Physical restraint</th>
<th>Protocol adherence</th>
<th>Turning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aloush (2018) Jordan</td>
<td>CS</td>
<td>N:P</td>
<td>171</td>
<td>171</td>
<td>15</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Bakhru (2016) Europe, US</td>
<td>CS</td>
<td>N:P</td>
<td>951</td>
<td>NR</td>
<td>951</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Benbenishty (2010) Europe</td>
<td>CS</td>
<td>N:P</td>
<td>NR</td>
<td>669</td>
<td>34</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Cho (2009) Korea</td>
<td>CS</td>
<td>N:P</td>
<td>1365</td>
<td>NR</td>
<td>65</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Dimick (2001) USA</td>
<td>RO</td>
<td>NNPR</td>
<td>NR</td>
<td>536</td>
<td>33</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Dodek (2012) Canada, France</td>
<td>RO</td>
<td>N:P</td>
<td>NR</td>
<td>30</td>
<td>2</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Goldhill (2008) UK</td>
<td>PO, L</td>
<td>N:P</td>
<td>NR</td>
<td>393</td>
<td>39</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Liu (2016) China</td>
<td>CS</td>
<td>N:P</td>
<td>1890</td>
<td>NR</td>
<td>134</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Roos-Blom (2020) Netherlands</td>
<td>CS</td>
<td>N:P</td>
<td>NR</td>
<td>8136</td>
<td>13</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
<tr>
<td>Rose (2011) Multiple</td>
<td>CS</td>
<td>N:P</td>
<td>NR</td>
<td>586</td>
<td>586</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>▲</td>
</tr>
</tbody>
</table>

Design: CS cross-sectional, RO retrospective observational, PO prospective observational, L longitudinal. Nurse staffing measure: N:P Nurse to Patient ratio, NNPR Night-time Nurse to Patient Ratio; \(P^{m}\) perception of adequate staffing, NR – Not Reported. Internal, External Validity: ++ Strong; + Moderate, - Weak. \(^{a}\) Professional collaboration: decisions not made by physicians independently of nurses, during titration of ventilator settings, weaning method, determination of extubation readiness and weaning failure (Rose, 2011). \(^{b}\) DFST: decisions to forgo life-sustaining therapies. \(^{c}\) Protocol adherence: adherence to Central Line-Associated Bloodstream Infection (CLABSI) prevention protocol (Aloush, 2018); \(^{i}\) frequency Richmond Agitation-Sedation Scale measurement is taken, \(^{ii}\) appropriateness of response to levels of sedation outside the target range (Dodek, 2011); Adherence to sepsis protocol (Kim, 2012). ▲ Higher staffing is significantly \((p<0.05)\) beneficial, △ Higher staffing is numerically beneficial. ▼ Higher staffing is significantly \((p<0.05)\) detrimental, ▼▼ Higher staffing is numerically detrimental. □ No evidence of a significant association \((p\geq0.05)\) and no figures given.
<table>
<thead>
<tr>
<th>First Author</th>
<th>Year</th>
<th>Country</th>
<th>Design</th>
<th>Nurse Staffing Measure</th>
<th>Number of RNs</th>
<th>Number of patients</th>
<th>Number of ICUs</th>
<th>Burnout</th>
<th>Job dissatisfaction</th>
<th>Intention to leave</th>
<th>Family satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azoulay (2001) France</td>
<td></td>
<td>France</td>
<td>PO</td>
<td>N:P</td>
<td>920</td>
<td>637</td>
<td>43</td>
<td>+</td>
<td>++</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Cho (2009) Korea</td>
<td></td>
<td>Korea</td>
<td>CS</td>
<td>N:P</td>
<td>1365</td>
<td>NR</td>
<td>65</td>
<td>++</td>
<td>++</td>
<td>▲ ▲ ▲</td>
<td>▲</td>
</tr>
<tr>
<td>Gerasimou-Angelidi (2014)</td>
<td></td>
<td>Greece</td>
<td>RO</td>
<td>NAS</td>
<td>161</td>
<td>NR</td>
<td>1</td>
<td>+</td>
<td>-</td>
<td>▲ ▲ ▲</td>
<td>▲</td>
</tr>
<tr>
<td>Johnson (1998) Canada</td>
<td></td>
<td>Canada</td>
<td>CS</td>
<td>N:P</td>
<td>115</td>
<td>99</td>
<td>1</td>
<td>-</td>
<td>+</td>
<td>▲</td>
<td>▲</td>
</tr>
<tr>
<td>Stricker (2009) Switzerland</td>
<td></td>
<td>Switzerland</td>
<td>CS</td>
<td>N:P</td>
<td>1321</td>
<td>NR</td>
<td>26</td>
<td>++</td>
<td>+</td>
<td>▲</td>
<td>▲</td>
</tr>
</tbody>
</table>

CS cross-sectional, RO retrospective observational, PO prospective observational. Nurse Staffing Measures: Ratio antecedents and consequents may have been calculated either way round (e.g., as N:P or P:N). N:P Nurse to Patient ratio. NAS Nurses activity score. NR Not Reported. ▲ Higher staffing is significantly (p<0.05) beneficial. ▲ Δ Higher staffing is numerically beneficial. ▲ ▲ Higher staffing is significantly (p<0.05) detrimental. ▲ ▲ ▲ Higher staffing is numerically detrimental. □ No evidence of a significant association (p≥0.05) and no figures given.
References


and factors that may influence staffing requirements. doi:URI:
http://eprints.soton.ac.uk/id/eprint/367526


