International Journal of Nursing Studies 134 (2022) 104311

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/locate/ns

Nurse staffing levels and patient outcomes: A systematic review of longitudinal studies

Chiara Dall'Ora [a,b,*], Christina Saville [a], Bruna Rubbo [a], Lesley Turner [a], Jeremy Jones [a], Peter Griffiths [a,b,c]

[a] School of Health Sciences, University of Southampton, Southampton, United Kingdom
[b] Applied Research Collaboration Wessex, University of Southampton, Southampton, United Kingdom
[c] Center for Health Outcomes and Policy Research, University of Pennsylvania, Philadelphia, United States of America

ARTICLE INFO

Article history:
Received 10 March 2022
Received in revised form 4 May 2022
Accepted 9 June 2022

Keywords:
Inpatients
Longitudinal studies
Nursing staff
Personnel staffing and scheduling
Workforce

ABSTRACT

Background: The contribution of registered nurses towards safe patient care has been demonstrated in many studies. However, most of the evidence linking staffing levels to outcomes is cross-sectional with intrinsic limitations including an inability to establish that presumed cause (staffing) precedes the effect. No reviews have summarised longitudinal studies considering nurse staffing and patient outcomes.
Objectives: To identify and assess the evidence for an association between nurse staffing levels, including the composition of the nursing team, and patient outcomes in acute care settings from longitudinal studies.
Methods: We undertook a systematic review of studies where the association between nurse staffing with patient outcomes was assessed in a longitudinal design. Studies with repeated cross-sectional analyses were excluded unless a difference-in-difference design was used. We searched Medline, CINAHL, Embase and the Cochrane Library up to February 2022. We used the ROBINS-I tool to assess risk of bias. We synthesised results in a tabular form and a narrative grouped by outcome.
Results: 27 papers were included. Studies were conducted in a variety of settings and populations, including adult general medical/surgical wards and adult and neonatal intensive care units. Staffing measures were operationalised in a variety of different ways, making direct comparisons between studies difficult and pooled estimates impossible. Most studies were either at serious (n = 12) or critical (n = 5) risk of bias, with only 3 studies at low risk of bias. Studies with the most risk of bias were judged as likely to underestimate the effect of higher registered nurse staffing. Findings are consistent with an overall picture of a beneficial effect from higher registered nurse staffing on preventing patient death. The evidence is less clear for other patient outcomes with a higher risk of bias, but in general the proposition that higher registered nurse staffing is likely to lead to better patient outcomes is supported. Evidence about the contribution of other nursing staff groups is unclear.
Conclusion: The causal relationship between low registered nurse staffing and mortality is plausible and these estimates of relationships from longitudinal studies provide further support. To address residual uncertainties, future studies should be conducted in more than one hospital and using standardised measures when reporting staffing levels.
Tweetable abstract: Having more registered nurses on hospital wards is causally linked to reduced mortality – new review shows there is little room for doubt @ora_dall @workforcesoton @turnel.

What is already known

- Many cross-sectional studies and reviews have demonstrated that higher registered nurse staffing levels are associated with better patient outcomes
- Longitudinal studies have the potential to establish a temporal link between presumed cause (variation in staffing levels) and effect (variation in patient outcomes)

What this paper adds

- Evidence from studies at low risk of bias confirmed a beneficial effect of increased registered nurse staffing levels on patient mortality
- Variation in the measurement of staffing inputs means that it remains difficult to estimate the effects of staffing increases

* Corresponding author at: School of Health Sciences, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, United Kingdom.
E-mail addresses: C.Dallora@soton.ac.uk (C. Dall'Ora), C.E.Saville@soton.ac.uk (C. Saville), B.Rubbo@soton.ac.uk (B. Rubbo), L.Y.Turner@soton.ac.uk (L. Turner), Jeremy.jones@soton.ac.uk (J. Jones), peter.griffiths@soton.ac.uk (P. Griffiths)
@ora_dall, @turnel, @workforcesoton.

https://doi.org/10.1016/j.ijnurstu.2022.104311
0020-7489/

- Evidence about the effect of staff other than registered nurses, including assistants, is unclear with mixed evidence showing both benefit and harm

1. Introduction

Nursing staff account for half of the health workforce worldwide, and the cost of providing inpatient hospital nursing is one of the main cost drivers for health systems (World Health Organization, 2016). Providing nurse staffing levels that match patient demand is key to deliver cost-effective health services. In the face of increasing financial pressures and budget constraints, registered nurses are sometimes viewed simply as a costly labour input, which can often be substituted with lower paid unregistered staff (Yakusheva et al., 2020). This has raised questions about the contribution of registered nurses and other nurses in the workforce in ensuring patients receive safe and high-quality care, including preventing deterioration, adverse outcomes such as infections and pressure ulcers, and death among hospital inpatients. Besides registered nurses, the nursing workforce comprises of unregistered nursing assistants, and licenced practical nurses/nursing associates, who access the profession by completing a shorter training programme and, therefore, have a reduced scope of practice compared to registered nurses.

When considering the associations between nurse staffing levels, the composition of the nursing workforce (i.e. skill mix) and patient outcomes in inpatient hospital settings, the breadth of the evidence is apparent, with hundreds of studies published to date, most of which support a conclusion that the higher the registered nurse staffing, the lower the rate of adverse patient outcomes, including death and infections (Griffiths et al., 2016). The volume of the evidence has led many to question whether more studies are needed because the implications for policy and practice are clear and the evidence definitive (Buchan et al., 2020; Beech et al., 2019). However, past reviews of this literature (Kane et al., 2007; Shekelle, 2013; Driscoll et al., 2018; Shin et al., 2019; Twigg et al., 2019) have noted the preponderance of cross-sectional studies. Such studies are unable to establish that the observed variation in staffing levels and skill mix between hospitals, typically measured at a hospital level average, corresponded to staffing that was experienced by the patients whose outcomes were measured, typically aggregated at the hospital level over a year.

The limitations of the evidence are such that for some commentators a causal interpretation of the relationship is still questionable, although a careful analysis applying epidemiological principles suggests that the body of evidence is indeed consistent with a cause and effect relationship (Griffiths et al., 2016; Kane et al., 2007). Nonetheless, the indirect associations reported in most studies means it is impossible to estimate the effect of change in staffing levels without bias.

In recent years, there has been a steady increase in studies using routinely collected data in healthcare, including studies that use electronic rostering systems and patient records to link patients to the staffing levels they are exposed to throughout their hospital stay (Musy et al., 2020; Haegdorens et al., 2019). Such studies have the potential to establish the staffing levels that individual patients have been exposed to *prior* to experiencing the outcome and so directly explore the effect of variation. In addition to removing many potential sources of bias associated with cross-sectional studies, longitudinal studies avoid the intrinsic limitation of cross-sectional studies, as the presumed cause can be shown to precede the outcome of interest. There is no summary of studies using longitudinal designs to explore the impact of staffing levels and skill mix on patient and organisational outcomes in inpatient hospital settings. Therefore, the aim of this systematic review is to summarise the effect of exposure to variation in nurse staffing levels, including skill mix, on subsequent patient outcomes in longitudinal studies.

2. Methods

We registered our review protocol in Prospero (PROSPERO 2020 CRD42020191798). We planned to include midwifery staffing in our review, as detailed in our review protocol, but since no studies on midwifery met the inclusion criteria, we framed the paper focusing on nurse staffing only.

2.1. Eligibility criteria

To be included, studies must estimate the effect on patients of exposure to varying levels or composition of nursing staff in a hospital inpatient unit (including intensive care units and general wards). Any measure of nurse staffing level or mixture of nursing staff was considered, including staff-to-patient ratios, staff hours per patient day, deviation in staffing from an established norm or reference (e.g. 'low staffing' relative to a defined standard), measured workload relative to available staffing, or relative mix of registered staff to other groups.

We included prospective, retrospective, cohort, case–control, randomised or quasi-randomised controlled trials, controlled before-and-after studies, interrupted time series, difference-in-difference, or panel studies. To be eligible, studies must establish a temporal link between nurse staffing and patient outcomes, such that measured variation in nurse staffing levels clearly occurred before the outcomes. Studies with repeated measures that used a cross-sectional analysis were excluded except where difference-in-difference designs were used to show an association between change in staffing and change in outcomes over time. We included studies where planned interventions were implemented and natural experiments where the effect of an exogenous 'shock' (e.g. changes in legislation or major policies designed to alter staffing levels) involving a change in staffing was studied. We did not specify a list of patient outcomes a priori, but we defined patient outcomes as any outcomes experienced by patients as opposed to staff, families, and the healthcare systems.

2.2. Study selection and data extraction

We searched Medline, CINAHL, Embase and the Cochrane Library. The complete search strategy can be found in Supplementary Material Table 1. Searches were undertaken up to February 2022.

One reviewer (CDO) undertook the first screening to remove duplicates and irrelevant studies. Potentially relevant papers were then further screened with a more detailed assessment of titles and abstracts. At this stage, all other reviewers assessed samples of 10 studies each to ensure that there was consistency, and to identify points of ambiguity and uncertainty in selection criteria. Full texts of studies that remained after this screening were retrieved and detailed assessment was made against the criteria. All full text papers were assessed by CDO and another reviewer. Disagreements were resolved by discussion to reach consensus among all reviewers. From each included study we extracted author(s), year; country and setting; sample size; measure(s) of staffing levels; outcomes and risk-adjustment; findings.

2.3. Risk of bias

We used the ROBINS-I tool (Sterne et al., 2016) to assess risk of bias because it supports detailed assessment of confounding, which is important as bias due to omitted variables occurs frequently in studies in this area (Griffiths et al., 2016). The ROBINS-I tool is designed to assess non-randomised studies and can be used for studies with a cohort design, in which individuals who have been exposed to variation in staffing levels are followed up over time (Sterne et al., 2016). Bias is defined as a tendency for study results to differ systematically from the results expected from hypothetical target randomised trial, conducted on the same participants and with no flaws in its conduct (Sterne et al., 2016). When using the tool to assess studies of natural variation in staffing we interpreted "intervention" as exposure of patients to different levels of staffing. The hypothetical target trial we identified was a randomised control trial where any patient, upon admission into a hospital unit, is randomly allocated to otherwise identical wards with different staffing levels or skill mix, or a trial in which staffing levels were varied at random (within constraints) throughout the patient's stay. For natural experiments the target trial was a cluster randomised controlled trial assessing the same intervention or policy change.

Using the ROBINS-I tool, the risk of bias due to different methodological aspects (called "domains") was assessed: confounding, selection of participants into the study, classification of interventions, deviation from intended interventions, missing data, measurement of outcomes, and selection of the reported results. A detailed explanation of how to grade each domain is available in the ROBINS-I guidance, where a table for each domain provides clear criteria to guide risk of bias judgements (Sterne et al., 2016). Risk of bias in each domain was graded as either low, moderate, serious, or critical. The domain grade with the highest risk of bias determined the overall risk of bias grade. According to the ROBINS-I guidance (Sterne et al., 2016), we excluded studies at critical risk of bias from our synthesis (Ambrosi et al., 2017; Jansson et al., 2019; Mark and Belyea, 2009; Palese et al., 2016; Twigg et al., 2016) but these studies' characteristics and detailed description of critical risk of bias are reported in Supplementary Material – Table 2. Where we identified a risk of bias we considered whether the mechanisms of bias were likely to be associated with an over- or under-estimate of nurse staffing effects, although in some cases the likely direction of bias was unclear.

Two reviewers independently assessed the risk of bias for each study, and where there was disagreement, this was resolved by collective discussion. As some reviewers were authors of included papers we ensured that at least one reviewer was not an author of the paper in each case and where possible allocated such papers to two non-authors.

2.4. Evidence synthesis

We were unable to identify groups of studies that used measures of staffing and outcomes that were sufficiently comparable to pool in a statistical meta-analysis. Therefore, we performed a narrative synthesis with results grouped by outcome. Where studies performed more than one analysis, we reported results at the analysis level rather than summarising them at the study level. Outcome measures were grouped for reporting if they were available in fifteen or more analyses. When less than fifteen analyses reported on the same outcome, we grouped outcomes in common themes.

3. Results

Our search yielded 4518 records, of which 946 were duplicates and 3280 were excluded based on title and abstract. We assessed the full text of the remaining 292 studies for eligibility. Of these, 27 published between 2003 and 2021 were included in the review (see Fig. 1). Most studies were single hospital (n = 15), but there were some notable exceptions. For example, McHugh et al. included 55 hospitals (McHugh et al., 2021), Hamilton et al. included 54 hospitals (Hamilton et al., 2007) and Mark and Belyea included 145 hospitals (Mark and Belyea, 2009). Patient samples ranged between 85 (Jansson et al., 2019) and 489,155 (McHugh et al., 2021). Twelve studies were conducted in Intensive Care Units (ICU); three in Acute Medical Units; twelve in a variety of inpatient wards, which could also include ICUs. Studies were conducted in the United States (n = 9), United Kingdom (n = 5), Canada (n = 4), Australia (n = 3), Switzerland (n = 3), Italy (n = 2) and Finland (n = 1). Nurse staffing levels were measured in a variety of ways, and these are summarised in Table 1. Some studies used multiple measures of staffing, for example both cumulative sum of nursing hours per patient day across all stay and for only part of a patient stay (i.e. first two or five days of patient stay). Descriptive information of studies including settings, designs, sample sizes, data sources, outcomes, risk adjustment and analysis level and type are displayed in Table 2.

The "exposure window" (i.e. the period of staffing measured prior to the outcome) ranged between 6 h (Al-Abdwani et al., 2018) and 30 days in studies with individual patient exposures (Mark and Belyea, 2009). In studies considering policy implementations, staffing and outcomes relationships were measured one year after implementation (McHugh et al., 2021; Twigg et al., 2016). Twelve studies considered staffing levels averaged or summed cumulatively for the whole patient stay, while six considered staffing in the early part of the patient stay only, ranging between the first day and the first seven days of the patients' stay (see supplementary material Table 3).

3.1. Risk of bias

Details of the risk of bias assessments by study are given in Supplementary Material Table 4a and b. Of the 27 studies, only three were assessed as at overall low risk of bias, seven were at moderate risk of bias, and twelve were classified as at serious risk of bias. Five studies were at critical risk of bias and so results are not reported below.

Eleven studies had serious risk of bias resulting from confounding. Studies lacked appropriate adjustment for risk of adverse outcome at baseline and/or failed to address time-varying confounding resulting from potential staffing increases as a response to worsening of a patients' condition. Bias in selection of participants into the study was either low or moderate, apart from one study of the implementation of staffing policies, which was classified as serious risk of bias due to the selection of units and hospitals which implemented staffing interventions and controls, which was not random. Bias in classification of interventions was found to be serious only in one study because the majority of studies used routinely collected data from reliable sources to determine staffing levels. One study had a serious risk of bias in the domain of deviation from interventions, as there was no evidence that co-interventions (temporary staffing levels) were balanced across intervention groups. Bias due to missing data was low or moderate in all studies apart from one, where authors report there are missing data, but do not quantify the magnitude. One study had a serious risk of bias arising from outcome measurement as procedures for monitoring the outcome were not reported; other studies had either low or moderate risk of bias. Six studies were potentially affected by bias in results reporting as they presented several subgroup analyses that were not justified in the methods and did not appear to be pre-planned.

In many studies, the direction of bias was unpredictable. However, in eight studies with a serious or critical risk of bias, lack of individual risk-adjustment would likely bias results in favour of lower staffing or attenuate any observed effects from higher staffing. Failure to include known predictors of the outcome in statistical models could have led to estimates favouring lower staffing levels because patients at low risk of experiencing a negative outcome have lower need and may consequently be exposed to lower staffing (He et al., 2013; Antonakis et al., 2010). On the other hand, in two studies failure to control for levels of other staff could lead to an overestimate of the benefits of higher nurse staffing levels because the nurse staffing measure was likely to correlate with staffing by other staff groups and these staff groups have an effect on outcomes. While some other studies had no direct control for other staff groups the effect of the omission was less predictable as the design meant that variation in nurse staffing was unlikely to be correlated with that of other staff groups or else there was evidence presented for a lack of correlation.

3.2. Mortality

Most analyses showed higher nurse staffing to be associated with reduced mortality. Ten studies reported 33 analyses on patient mortality.

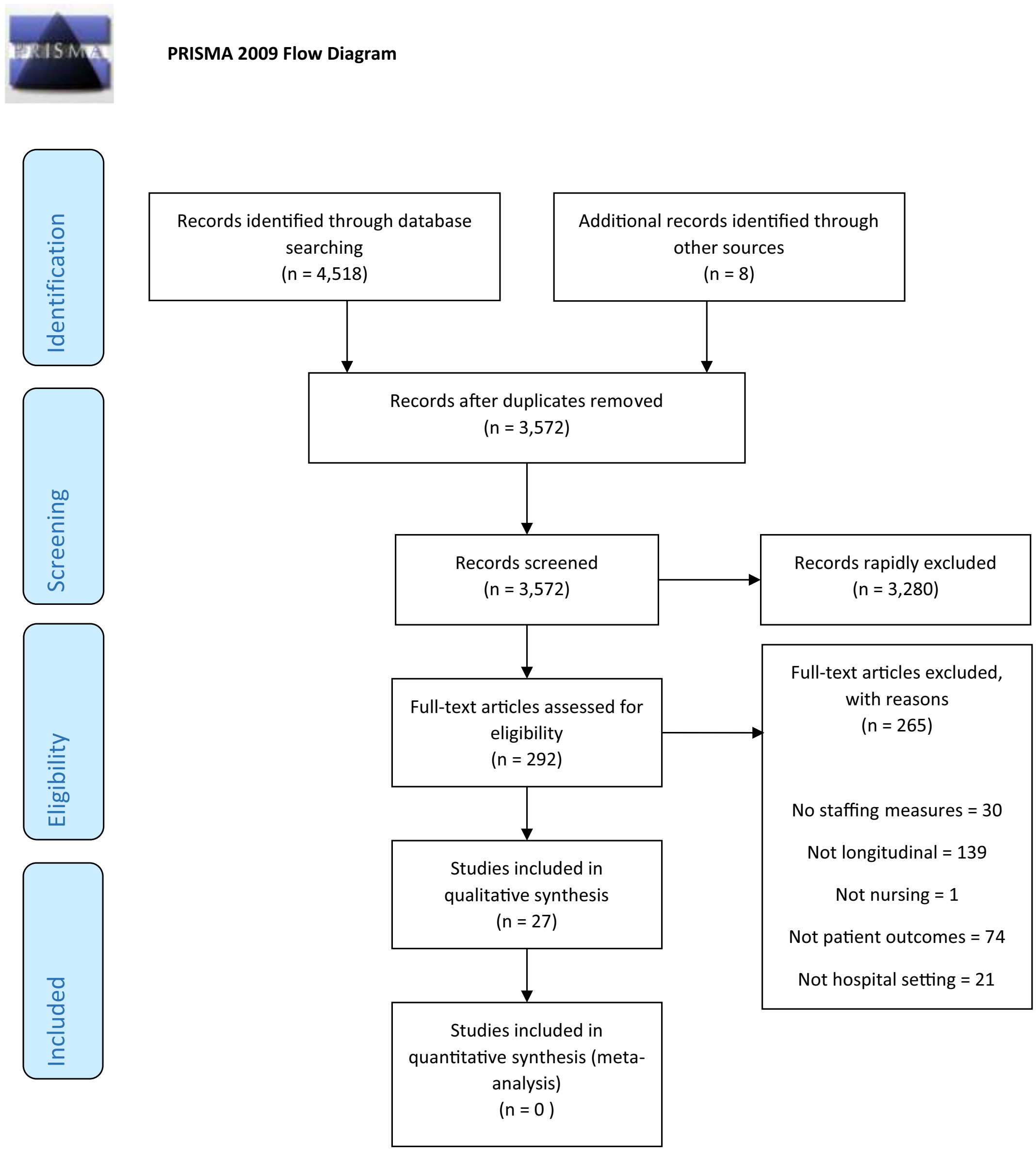

Fig. 1. PRISMA flow diagram adapted from Page et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: https://doi.org/10.1136/bmj.n71, distributed under the terms of the Creative Commons Attribution License

We report estimates for associations between registered nurse staffing and mortality in Table 3. Effect sizes were, typically, small. Exposure to low staffing was measured with different thresholds and different exposure windows. The diversity of staffing measures and exposure windows makes any meaningful comparison or synthesis impossible. Focusing on registered nurses staffing levels (19 analyses in total), most analyses showed higher staffing to be associated with reduced mortality although there was some inconsistency. Studies favouring higher staffing included studies with larger samples in terms of number of wards and patients at low/moderate risk of bias. Twelve analyses from seven studies, including three at low risk of bias, found that higher registered nurse staffing levels were associated with a reduced risk of in-hospital mortality in adult general patient populations (Beltempo et al., 2018; Griffiths et al., 2019; Needleman et al., 2011; Needleman et al., 2020; Rochefort et al., 2020; Fogg et al., 2021; Musy et al., 2021). Six analyses from four studies reported associations between registered nurse staffing levels and patient mortality that were not statistically significant (Hamilton et al., 2007; Rochefort et al., 2020; Needleman et al., 2020; Callaghan et al., 2003), although coefficients, when presented, favoured higher registered nurse staffing. A single analysis from a study with small sample size at serious risk of bias found that higher registered nurse staffing levels were associated with higher risk of mortality among 692 patients in a single ICU (Callaghan et al., 2003).

Analyses looking at the effect of nursing assistants or licenced practical nurses did so either focusing on levels of these staff or by considering skill-mix in the team. Findings were mixed and provide an inconsistent picture with contrasting results. Effect size estimates are reported in Table 3. Three analyses (two studies) reported increased mortality when patients were exposed to low nursing assistant staffing (Griffiths et al., 2019; Needleman et al., 2020), although Griffiths et al. found a non-linear relationship so nursing assistant staffing above the norm was also associated with increased risk of mortality (Griffiths et al., 2019). A non-linear effect was also apparent in Musy et al.'s study, whereby both high and low licenced practical nurse staffing levels were associated with lower mortality, although only the adverse effect of high licenced practical nurse staffing levels was statistically significant (Musy et al., 2021). Two analyses from two studies found higher mortality when patients were exposed to higher nursing assistant staffing levels, although these results were not statistically significant (Fogg et al., 2021; Griffiths et al., 2019). Needleman and colleagues also combined registered nurse and nursing assistant staffing levels in two analyses and found patients exposed to lower staffing levels from both groups were more likely to die (Needleman et al., 2020). Skill-mix was explored in one study only, and it was not associated with patient mortality (Rochefort et al., 2020).

Three analyses from one study at serious risk of bias in one hospital with 138,133 patients found that exposure to higher levels of temporary registered nurse and nursing assistant staffing were associated with higher patient mortality. For registered nurses this effect was statistically significant only when a high number of registered nurse hours, equivalent to between one third and one half of the average ward staffing complement, were from temporary nurses (HR = 1.12; 95% CI = 1.03–1.21) (Dall'Ora et al., 2020).

3.3. Infections

Sixteen analyses from seven studies examined the impact of staffing levels on infections, including central venous catheter associated bloodstream infections, early and late onset ventilator-associated pneumonia, and healthcare-associated infections. Overall, there was inconclusive evidence of the effect of nurse staffing levels on infections, although results generally favoured higher registered nurse staffing levels. Effect estimates are reported in Table 4.

Eight analyses from seven studies found no statistically significant associations with registered nurse staffing levels, although most point estimates, when available, showed a protective effect of higher registered nurse staffing (Alonso-Echanove et al., 2003; Beltempo et al., 2017; Hugonnet et al., 2007a; Hugonnet et al., 2007b; Shang et al., 2019). The exception was a single analysis in one of the two neonatal ICUs where higher RN hours per patient day were associated with higher hazard of infection in one unit and lower hazard in the other (not statistically significant) (Cimiotti, 2004; Cimiotti et al., 2006) In adult intensive care settings (Hugonnet et al., 2007a; Hugonnet et al., 2007b) and adult hospital settings (Shang et al., 2019) higher registered nurse staffing was associated with a statistically significant reduction in infections.

One analysis found that higher nursing assistant staffing levels were associated with lower risk of healthcare-associated infections (Shang et al., 2019) while two analyses found no statistically significant relationships between nursing assistant staffing and infections (point estimates not reported) (Alonso-Echanove et al., 2003; Shang et al., 2019). One analysis found that patients exposed to days with higher levels of temporary nurse staffing were more likely to experience central venous catheter-associated bloodstream infections (Alonso-Echanove et al., 2003).

3.4. Other outcomes

In a single site study of 138,133 patients, Griffiths and colleagues examined the effect of staffing levels on a composite outcome of adverse events, including death, cardiac arrest or unplanned ICU admission. Exposure to higher than the mean registered nurse hours per patient day in the first five days of a hospital admission was associated with a reduced hazard of experiencing adverse events (HR = 0.98; 95% CI = 0.96–0.99). Results for nursing assistant hours per patient day were in the opposite direction, although not statistically significant (HR = 1.01; 95% CI = 0.99–1.02) (Griffiths et al., 2018b).

Patrician and colleagues examined the effect of staffing levels and skill mix on hospital-acquired pressure injuries in 13 hospitals and 1643 patients through 12 distinct analyses, exploring various exposure windows and skill mix configurations. The majority (10/12) of these analyses found no statistically significant associations, but two analyses found that a skill mix richer in licenced practical nurses over the previous three days (HR = 0.27 (no 95% CI reported)) and one week (HR = 0.56

Table 1
Frequency of staffing level and skill mix measures used in included studies.

Staffing measure	Number of studies	Studies
Staffing levels		
Studies considering absolute staffing levels		
Patient-to-nurse ratio	9	(Callaghan et al., 2003, Rochefort et al., 2020, Hugonnet et al., 2007a, Hugonnet et al., 2007b, Alonso-Echanove et al., 2003, McHugh et al., 2021, Jansson et al., 2019, Mark and Belyea, 2009)
Nurse Hours per Patient Day	4	(Cimiotti, 2004, Cimiotti et al., 2006, Tschannen, 2005, Palese et al., 2016)
Nurse Hours per Patient per 12-hour shift	1	(Needleman et al., 2020)
Nurse Hours per shift	1	(Rochefort et al., 2020)
Nurses per shift	1	(Hamilton et al., 2007)
Nurses prior to the event	1	(Al-Abdwani et al., 2018)
Studies considering relative to a defined standard		
Days with staffing below unit mean	2	(Griffiths et al., 2018b, Griffiths et al., 2019)
Shifts with low Nurse Hours per Patient per Shift ("low" was defined differently in each study)	2	(Needleman et al., 2020, Shang et al., 2019)
Nurse hours per patient day relative to the mean	2	(Fogg et al., 2021, Griffiths et al., 2019)
Nurses available divided by the number of recommended nurses per shift	2	(Hamilton et al., 2007, Beltempo et al., 2017)
RN minutes per day at weekend vs weekday	1	(Ambrosi et al., 2017)
Nurse Hours worked divided by the total number of required hours of care based on patient dependency categories	1	(Beltempo et al., 2018)
50% above/below median number of RN expected	1	(Musy et al., 2021)
Nurse Hours per Patient Day ≥8 h below requirement	1	(Needleman et al., 2011)
Shifts with Nurse Hours lower than expected by ≥8 h	1	(Rochefort et al., 2020)
Composition/staff mix of the nursing team		
Skill mix (RNs/RNs + LPNs + NAs)	2	(Patrician et al., 2017, Mark and Belyea, 2009)
Skill mix (NA/RNs + LPN + NAs)	1	(Patrician et al., 2017)
Skill mix (LPN/RNs + LPN + NAs)	1	(Patrician et al., 2017)
Skill mix per shift (NonRN/RN + NonRN)	1	(Rochefort et al., 2020)
Being cared for by float nurses for >60% of the time	1	(Alonso-Echanove et al., 2003)
Days with additional Temporary Nurse Hours per Patient Day	1	(Dall'Ora et al., 2020)
Specialist nurses available divided by the number of recommended specialist nurses per shift	1	(Hamilton et al., 2007)
Ratio of ICU trained to trainee nurses	1	(Hugonnet et al., 2007a)

RN: Registered Nurse; NA: Nursing Assistant; LPN: Licenced Practical Nurse;

Table 2
Settings, designs, sample sizes, data sources, outcomes, risk adjustment and analysis level & type for included studies.

Author(s), year	Country & setting	Design & Data collection period	Sample size	Data source (routine, survey)	Outcome	Risk-adjustment	Analysis level & type
Al-Abdwani et al., 2018	Canada; Paediatric ICU and Cardiac ICU	Retrospective observational; 01/2004–01/2015	Hospitals = 1; ICUs = 2; Nurses = N/A; Patients = N/A; Patient admissions = 11,310	**Outcomes**: patient safety database + electronic health records **Nurse staffing**: electronic health records	Use of positive pressure ventilation within 24 h of unplanned extubation.	Age, gender, intubation reason, intubation duration, ICU day of the event, secretions, previous intubations, shared room, day time, weekday, ICU census, ICU vent census, Nasal intubation, cuffed endotracheal tube, endotracheal tube taped, high endotracheal tube, re-tape endotracheal tube, chest radiograph, clinical activities, conventional ventilations, nurse ICU experience, respiratory therapists.	Patient-day level; Logistic regression
Alonso-Echanove et al., 2003	USA; Medical ICU; Surgical ICU	Prospective observational; 12/1997–12/1999	Hospitals = 6; ICUs = 8; Patients = 4535; Nurses = N/A	**All**: Daily survey. Data were collected by infection control practitioners and sent to the CDC.	CVC-associated blood stream infections.	CVC in a patient unarousable for more than 70% of the CVC-days, an interaction term for TPN and non-antimicrobial-impregnated CVCs, patient not receiving antibiotics within 48 h of the CVC insertion, the use of peripherally inserted central catheters, patient in age category 45–55 yrs. old	Patient level; Cox proportional hazard model
Beltempo et al., 2017	Canada; Neonatal ICU	Case control; 04/2011–03/2013	Hospitals = 1; Wards = 1; Nurses = 165; Patients = 2236	**Outcomes**: Clinical database **Staffing levels**: Administrative data feeding into payroll	Healthcare associated infection	Nurse overtime, unit occupancy, n° patients with central lines, no of acutely ill patients and the no of daily admissions	Ward-day level; Logistic regression
Beltempo et al., 2018	Canada; Neonatal ICU	Cohort study; 03/2011–06/2014	Hospitals = 1; Wards = 1; Nurses = 165; patients = 257	**Outcomes**: local Canadian Neonatal Network database. **Nurse staffing**: administrative software used for scheduling and payroll	Mortality (all causes) or major morbidity.	Gestational age, sex, small for gestational age status, outborn, SNAP-II score more than 20, mode of delivery, overtime and occupancy for the first day of admission, first 7 days and NICU hospitalizations	Patient level; Multi-variable log-binomial regression
Callaghan et al., 2003	Australia; Neonatal ICU	Cohort study; 01/1996–12/1999	Hospitals = 1; Wards = 1; Patients = 692 =	**All**: Routinely collected data from electronic systems	In-hospital mortality	Initial risk using the Clinical Risk Index for Babies (CRIB) score, and dependency areas (ICU, high dependency, medium dependency or recovery)	Patient level; Logistic regression
Cimiotti et al., 2004; Cimiotti et al., 2006	USA; Neonatal ICU	Cohort study; 03/2001–02/2003	Hospitals = 1; Wards = 2; Patients = 2675	**Outcome**: Clinical data from patients' records. **Nurse staffing**: Nurse Staffing Office of each ward	Healthcare associated bloodstream infection	Birth weight, catheter use, major surgery, total parenteral nutrition, hand hygiene product used by staff	Patient level; Cox proportional hazards regression
Dall'Ora et al., 2020	UK; General hospital, medical/surgical	Retrospective; 04/2012–04/2015	Hospitals = 1; Wards = 32; Patients = 138,133; RNs = 1244; HCAs = 700; Nursing shifts = 761,946	**Outcome**: Patient Administration System **Nurse staffing**: Electronic rostering system + bank & agency database	In-hospital mortality	National Early Warning Score on admission; Summary Hospital Mortality Indicator risk score; emergency vs. elective admissions; admissions per RN > 125% of unit mean; admissions per NA > 125% of unit mean; RN staffing below unit mean; NA staffing below unit mean.	Patient level; hierarchical mixed-effects survival
Fogg et al., 2021	UK; General hospital, medical/surgical	Retrospective; 01/2014–04/2015	Hospitals = 1; Wards = 32; Patients = 9643	**Outcome**: Patient Administration System **Nurse staffing**: Electronic rostering system + bank & agency database	1) Death in hospital or within 30 days of dischargeUnplanned readmission to hospital within 30 days of discharge	Age; cognitive impairment; primary diagnosis group, based on Summary Hospital Mortality Indicator (SHMI) Clinical Coding System (CCS) bundles (original diagnoses use International Classification of Disease 10); Charlson co-morbidity index (CCI); NEWS value at admission; MUST score at admission; Route of admission; Ward transfers; Discharge specialty	Patient level; hierarchical mixed-effects survival
Griffiths et al., 2018a, 2018b	UK; General hospital, medical/surgical	Retrospective; 04/2012–04/2015	Hospitals = 1; Wards = 32; Patients = 138,133; RNs = 1244; HCAs = 700; Nursing shifts = 761,946	**Outcome**: cardiac arrest audit and ICU WardWatcher; Length of stay: Patient Administration System **Nurse staffing**: Electronic rostering system + bank & agency database	1) Adverse events (death, cardiac arrest or unplanned ICU admission)Length of stay in the hospital	NEWS on admission; SHMI risk score; age; Charlson co-morbidity index (CCI); primary diagnosis group; Type of admission (elective vs emergency); Admissions per RN of > 125% of ward mean; Admissions per HCA of > 125% of ward mean	Patient level; hierarchical mixed-effects survival
Griffiths et al., 2019	UK; General hospital,	Retrospective; 04/2012–04/2015	Hospitals = 1; Wards = 32; Patients = 138,133;	**Outcome**: Patient Administration System	In-hospital mortality	National Early Warning Score on admission; Summary Hospital Mortality Indicator risk score; emergency vs.	Patient level; hierarchical

	medical/surgical		RNs = 1244; HCAs = 700; Nursing shifts = 761,946	**Nurse staffing**: Electronic rostering system + bank & agency database		elective admissions; admissions per RN > 125% of unit mean; admissions per NA > 125% of unit mean; RN staffing below unit mean; NA staffing below unit mean	mixed-effects survival
Hamilton et al., 2007	UK; Neonatal ICU	Cohort study; 01/1998–04/1999	Hospitals = 54 Wards = 54; Patients = 2585	**Outcome**: recorded by research nurses **Nurse staffing**: recorded by research nurses twice daily	Death before discharge and planned deaths at home (excluded lethal malformations & deaths post specialist surgery)	Gestational age, gender, birth weight and mode of delivery, diagnostic category, maternal treatment with antenatal steroids, admission temperature, most extreme partial pressure of carbon dioxide (PaCO2), mean appropriate fraction of inspired oxygen (FiO2), and lowest base excess	Patient level; Logistic regression
Hugonnet et al., 2007a	Switzerland; Medical ICU	Cohort; 01/1999–12/2002	Hospitals = 1; Wards = 1; Patients = 1883; Patient-days = 10,637	**Outcome**: not specified **Staffing levels**: not specified	Healthcare associated infection	Presence of CVC, Ventilation, Urinary catheter, Antibiotics	Ward-day; Poisson regression
Hugonnet et al., 2007b	Switzerland; Medical ICU	Cohort; 01/1999–12/2002	Hospitals = 1; Wards = 1; patients = 2470	**Outcome**: not specified **Staffing levels**: not specified	Ventilator-Associated pneumonia; early-onset: 1–5 days after intubation; late-onset: after day 6)	For early-onset VAP adjusted for: Charlson score (comorbity index), having a CVC, being on a therapeutic antibiotic. For late-onset VAP: nursing acuity severity score, pulmonary disease, having a CVC, a peripheral venous line, and receiving therapeutic antibiotic	Patient level Cox regression analysis
McHugh et al., 2021	Australia, acute care hospitals	Prospective panel study; 2016 vs 2018	Hospitals = 55; nurses = 17,010; Patients = 231,902 at baseline, 257,253 post-implementation	**Nurse staffing**: Nurse survey **Outcomes**: Queensland Hospital Admitted Patient data and death records	1) 30-day mortality 2) 7-day readmissionsLength of stay	1) 2) 3): Hospital size; 1) and 2): 17 indicators from the Charlson Comorbidity Index, sex, age, and DRG	Hospital level; 1)&2) multilevel logistic regression; 3) zero-truncated negative binomial regression
Musy et al., 2021	Switzerland; Adult hospital	Observational study; 2015–2017	Hospital = 1; Wards =55; Patients = 79,893; RNs = 3646; LPNs = 438	**Outcome**: medical discharge data **Nurse staffing**: tacs® nurse staffing system	Inpatient mortality	Age; gender; residence before hospital admission; type of admission; decision-maker regarding admission; stay in Intensive Care; stay in emergency; LOS; n° of transfers; no of ICD-10-GM diagnoses	Patient level; Logistic regression
Needleman et al., 2011	USA; Tertiary academic medical centre	Observational study; 2003–2006	Hospital = 1; Wards = 43 Admissions = 197,961; Shifts = 176,696	**All**: Electronic data systems of the medical centre, electronic discharge abstracts	In-hospital mortality.	Each patient was assigned a predicted in-hospital mortality value based on the patient's diagnosis-related group (DRG).	Patient level. Cox proportional hazard
Needleman et al., 2020	USA; Academic health system	Observational; 2007–2012	Hospitals = 3 (2 tertiary and 1 community); Admissions: 78,303	**Outcome**: discharge summary. **Nurse staffing**: hospital payroll system.	In-hospital mortality	Age, gender, Charlson comorbidity score and month/year	Patient level. Cox proportional hazards
Patrician et al., 2017	USA; medical-surgical, stepdown and critical care units of military hospitals	Secondary longitudinal; 2003–2006	Hospitals = 13; Wards = 56 Patients = 1643; Shifts = 13,590	**All**: prevalence surveys collected annually from the Military Nursing Outcomes Database (MilNOD) project	Hospital-acquired pressure injury (HAPI) development (stage 2 or greater)	Age, Braden score, albumin, BUN, creatinine, unit day level, census, turnover and patient acuity	Patient level. Cox proportional hazards and generalised estimating equations
Rochefort et al., 2020	Canada; Adult medical, surgical, and ICU	Observational; 2010–2017	Hospitals = 1; Wards = 32; Patients = 146,349	**Outcome**: Patient administration data from discharge abstracts **Nurse staffing**: Payroll data	In-hospital mortality	Age, gender; Charlson Comorbidity Index; severity of illness on admission; Hospital admission type; unit of admission; year and month of admission; type of unit; timing of shift; current unit occupancy; weekend/statutory holiday; cumulative proportion of shifts spent in ICU; and the square of that cumulative proportion.	Patient level. Multivariable Cox proportional hazards regression
Shang et al., 2019	US; tertiary/quaternary & community hospitals	Observational; 2007–2012	Hospitals = 3; Wards = 34; Patients = 100,264; shifts = 66,871	**Outcome**: patient administrative system, clinical data warehouse, electronic health record system **Nurse staffing**: hospital payroll	Healthcare Associated Infections (urinary tract, bloodstream and pneumonia)	Patient individual factors; unit type, year, and patient turnover discharges	Patient level; Cox proportional-hazards regression
Tschannen, 2005	US, inpatient medical and surgical wards	Prospective; 03/2004–04/2004	Hospitals = 2; wards = 4; nurses = 135; patients = 310	**Outcome**: discharge data **Nurse staffing**: payroll records and census log for nurse staffing	1) Length of stay in the unitDeviation from expected LOS: expected length of stay was determined using Diagnostic Related Groups.	Gender, age, primary diagnosis (from DRG), Roman-Charlson Comorbidty index for co-morbidities; admission type (medical vs surgical); admission source (A&E, ICU, post-op); unit of admission	Patient level; regression analysis

Table 3
Mortality effect estimated for studies reporting variation in registered nurse staffing levels.

Authors, year	Staffing measure	Staffing group	Exposure window	Outcome	Result	Risk of bias
Effect of lower RN staffing						
Callaghan et al., 2003	Patient to nurse ratio of 1.71–1.97	RN	Averaged across first 72 h of stay	In-hospital mortality	Compared to exposure to ratio 1.16–1.58: OR = 0.18; 95% CI = 0.06–0.5*	Serious
Callaghan et al., 2003	Patient to nurse ratio of 1.59–1.70	RN	Averaged across first 72 h of stay	In-hospital mortality	Compared to exposure to ratio 1.16–1.58: OR = 0.84; 95% CI = 0.42–1.66	Serious
Griffiths et al., 2019	Days with staffing below the unit mean	RN	Cumulative sum of days over the first 5 days of stay	In-hospital mortality	HR = 1.03; 95% CI = 1.01–1.06*	Low
Needleman et al., 2011	Nurse Hours per Patient Day ≥8 h below requirement	RN	Cumulative sum across all stay	In-hospital mortality	HR = 1.02; 95% CI = 1.01–1.03*	Moderate
Needleman et al., 2011	Nurse Hours per Patient Day ≥8 h below requirement	RN	Cumulative sum of first five days of patient stay	In-hospital mortality	HR = 1.03; 95% CI = 1.02–1.05*	Moderate
Needleman et al., 2011	Nurse Hours per Patient Day ≥8 h below requirement	RN	Cumulative sum of 2 shifts prior to death	In-hospital mortality	HR = 1.05; 95% CI = 1.02–1.07*	Moderate
Needleman et al., 2020	Nurse hours per patient per 12-h shift	RN	Any shift with staffing below 75% of annual median unit	In-hospital mortality	HR = 1.02; 95% CI = 1.00–1.05*	Moderate
Needleman et al., 2020	Number of shifts with low nurse hours per patient per shift	RN	Cumulative sum over the 2nd to 5th day of stay	In-hospital mortality	HR = 1.04; 95% CI = 0.99; 1.10	Moderate
Rochefort et al., 2020	Shifts with nurse hours lower than expected by ≥8 h	RN	Cumulative sum across stay	In-hospital mortality	HR = 1.01; 95% CI = 1.00–1.01*	Low
Effect of higher RN staffing						
Beltempo et al., 2018	Total number of nurse hours worked divided by the total number of required hours of care based on patient dependency categories	RN	First day of NICU stay	Mortality (all causes) or major morbidity	RR = 0.83; 95% CI = 0.77–0.90*	Moderate
Beltempo et al., 2018	Number of RN hours worked divided by the number of required hours of care based on patient dependency categories	RN	Averaged across first 7 days	Mortality (all causes) or major morbidity	RR = 0.82; 95% CI: 0.76–0.89*	Moderate
Beltempo et al., 2018	Total number of nurse hours worked divided by the total number of required hours of care based on patient dependency categories	RN	Averaged across patient stay	Mortality (all causes) or major morbidity	RR = 0.81; 95% CI = 0.74–0.90*	Moderate
Fogg et al., 2021	RN hours per patient day relative to the mean (with non-linear terms)	RN	Averaged across patient stay relative to the mean	In-hospital mortality and within 30 days of discharge	Mean vs mean + 0.5 Registered Nurse Hours per Patient Day: OR = 0.90; 95% CI = 0.84–0.97*	Moderate
Griffiths et al., 2019	RN hours per patient day relative to the mean	RN	Cumulative sum over first 5 days of stay	In-hospital mortality	HR = 0.97; 95% CI = 0.94–0.99*	Low
Hamilton et al., 2007	Number of nurses per shift	RN	Averaged across patient stay	Death before discharge and planned deaths at home	No estimates - reported as nonsignificant	Moderate
Hamilton et al., 2007	Number of nurses available divided by the number of recommended nurses per shift	RN	Averaged across patient stay	Death before discharge and planned deaths at home	No estimates - reported as nonsignificant	Moderate
Musy et al., 2021	High/low staffing (50% above/below median number of RN expected given ward, patient activity & shift type)	RN	Cumulative count of shifts with high numbers across all stay	In-hospital Mortality	Higher staffing: OR = 0.91*; 95% CI = 0.89–0.93	Low
Rochefort et al., 2020	Nurse/patient ratio	RN	Cumulative sum across stay	In-hospital mortality	No estimates - reported as nonsignificant	Low
Rochefort et al., 2020	Nurse hours per shift	RN	Cumulative sum across stay	In-hospital mortality	No estimates - reported as nonsignificant	Low
Effect of skill mix & higher total nurse staffing						
Hamilton et al., 2007	Number of specialist nurses available divided by the number of recommended specialist nurses per shift	RN	Averaged across patient stay	Death before discharge and planned deaths at home	OR = 0.63; 95% CI = 0.42–0.96*	Hamilton et al., 2007
Needleman et al., 2020	Number of shifts with low nurse hours per patient per shift	RN + NA	Any shift with staffing below 75% of annual median unit	In-hospital mortality	HR = 1.02; 95% CI = 1.00–1.04*	Needleman et al., 2020
Needleman et al., 2020	Number of shifts with low nurse hours per patient per shift	RN + NA	Cumulative sum over the 2nd to 5th day of stay	In-hospital mortality	HR = 1.13; 95% CI = 1.08–1.18*	Needleman et al., 2020
Rochefort et al., 2020	Skill mix per shift (NonRN/RN + NonRN)	RN, NonRN	Cumulative sum across stay	In-hospital mortality	HR = 1.00; 95% CI = 0.99–1.01	Rochefort et al., 2020
Effect of lower assistant staffing						
Griffiths et al., 2019	Days with staffing below the unit mean	NA	Cumulative sum of	In-hospital	HR = 1.04; 95% CI = 1.02–1.07*	Low

Table 3 (*continued*)

Authors, year	Staffing measure	Staffing group	Exposure window	Outcome	Result	Risk of bias
			days over the first 5 days of stay	mortality		
Musy et al., 2021	High/low staffing (50% above/below median number of RN expected given ward, patient activity & shift type)	LPN	Cumulative count of shifts with high numbers across all stay	In-hospital mortality	Lower staffing: OR = 0.99; 95% CI = 0.96–1.01	Low
Needleman et al., 2020	Nurse hours per patient per 12-h shift	NA	Cumulative sum over the 2nd to 5th day of stay	In-hospital mortality	HR = 1.03; 95% CI = 1.00–1.05*	Moderate
Needleman et al., 2020	Nurse hours per patient per 12-h shift	NA	Any shift with staffing below 75% of annual median unit	In-hospital mortality	HR = 1.03; 95% CI = 1.01–1.04*	Moderate
Effect of higher assistant staffing						
Fogg et al., 2021	Hours per patient day relative to the mean (with non-linear terms)	NA	Averaged across patient stay relative to the mean	In-hospital mortality and within 30 days of discharge	Additional NA Hour: OR = 1.12, 95% CI = 0.91–1.39	Moderate
Griffiths et al., 2019	Hours per patient day relative to the mean	NA	Cumulative sum of days over the first 5 days of stay	In-hospital mortality	HR = 1.01; 95% CI = 0.98–1.04	Low
Musy et al., 2021	High/low staffing (50% above/below median number of RN expected given ward, patient activity & shift type)	LPN	Cumulative count of shifts with high numbers across all stay	Mortality	Higher staffing: OR = 0.97; 95% CI = 0.96–0.99*	Low

* Statistically significant at <0.05; OR = Odds Ratio; HR = Hazard Ratio; RR = Risk Ratio; CI = Confidence Interval; RN = Registered Nurse; NA = Nursing Assistant; LPN = Licenced Practical Nurse.

(no 95% CI reported)) reduced the hazard of hospital-acquired pressure injuries (Patrician et al., 2017).

Griffiths and colleagues found that patients' length of stay in hospital was reduced by a mean of 0.23 days for each additional registered nurse hours per patient day that a patient experienced (Gamma coefficient = −0.23; 95% CI = −0.30 - -0.16), while there was a small but statistically significant increase in stay for each additional nursing assistant hours per day a patient was exposed to throughout their stay (Gamma coefficient = 0.076; 95% CI = 0.03–0.13) (Griffiths et al., 2018b). Tschannen found that being exposed to higher nursing hours per patient day was associated with shorter than expected stays (relative to Diagnoses Related Groups based norms) (B = 2.481, SE = 1.0), but not overall average length of stay (B = 0.43, SE = 0.01) (Tschannen, 2005). A single site study of 9643 patients over the age of 75 who received a cognitive screening, found that patients exposed to an additional 0.5 registered nurse hours per patient day had a reduced risk of readmission, although this was not statistically significant (OR = 0.94; 95% CI = 0.82–1.06) (Fogg et al., 2021).

One study in two ICUs (one paediatric and one cardiac) and 11,310 admissions found that patients were more likely to receive positive pressure ventilation following unplanned extubation when there were more registered nurses in the preceding 6 h (OR = 1.53; 95% CI: 1.11–2.12) (Al-Abdwani et al., 2018).

3.5. Staffing policies

One study examining the effect of staffing policy interventions was included in our review (McHugh et al., 2021). It was a prospective panel study compared patient outcomes in hospitals that implemented a minimum nurse-to-patient ratio policy. Twenty-seven hospitals were subject to the staffing policy and 28 were not. Reducing workloads by one patient per nurse was associated with a decrease in 30-day mortality (OR = 0.93; 95% CI = 0.86–0.99), 7-day readmissions (OR = 0.93; 95% CI = 0.89–0.97) and length of stay (OR = 0.97; 95% CI = 0.94–0.99).

4. Discussion

This is the first systematic review of the effect of nurse staffing levels on patient outcomes focusing on longitudinal studies. These studies can demonstrate a temporal link between exposure to staffing levels and outcomes and thus overcome an intrinsic limitation of the majority of research in this field, which is cross-sectional. While studies were conducted across diverse samples, using different staffing measures and exposure windows, findings are consistent with higher registered nurse staffing reducing the risk of patient death. On the other hand, evidence about the positive contribution of other staff groups and changes to the skill mix of the nursing team is much more mixed although there is some evidence of harm linked to high levels of assistant staffing and temporary staff.

This evidence is consistent with conclusions based primarily on cross-sectional evidence (Griffiths et al., 2016), but when considering criteria to establish a causal relationship, there is a fundamental requirement that cause precedes effect (Hill, 1965; Rothman and Greenland, 2005). Although the evidence reviewed here remains observational, if relevant confounders are controlled for in the analyses, the parameter estimates for exposure can potentially be causally interpreted (Hernán, 2018), although some risk of bias remains. In addition, the causal pathway from low staffing to mortality has been theorised and empirically demonstrated, in part mediated by a failure to observe and mobilise response to deterioration (Smith et al., 2020; Griffiths et al., 2018b; Redfern et al., 2019; Ball et al., 2018).

In our review, studies at higher risk of bias were more likely to lead to an underestimation of the effect of higher nurse staffing than to overestimation. While in theory randomised controlled trials are a stronger study design with superior ability to determine a causal relationship (Higgins et al., 2021), it is not likely to be feasible under many circumstances, nor a guarantee of better quality if the study has risk of bias in other domains or is intrinsically limited by (for example) small sample size or randomisation of a small number of clusters. The three studies at low risk of bias in this review all concluded that exposure to higher staffing levels was associated with lower patient mortality. Studies at low risk of bias are able to provide effect estimates comparable to a well-performed randomised trial.

The diverse exposure windows, diverse settings and patient populations, and long span in which studies were published (i.e. 2003–2021) make it difficult to come to firm conclusions about the size of effects beyond the estimates provided by individual studies. The effect sizes observed for mortality are typically small but the large populations

Table 4
Infections effect estimated for analyses of staffing levels.

Study	Staffing measure	Staff group	Exposure window	Outcome	Result
Higher staffing					
Hugonnet et al., 2007a	Nurse/patient ratio	RN	Averaged across 4 days prior to infection	Healthcare associated infection	IRR = 0.69; 95% CI = 0.50–0.95*
Hugonnet et al., 2007a	Ratio of ICU trained to trainee nurses	RN	Averaged across 4 days prior to infection	Healthcare associated infection	IRR = 0.96; 95% CI = 0.82–1.13
Cimiotti, 2004; Cimiotti et al., 2006	Nurse Hours per Patient Day	RN	Averaged from 48 to 144 h before infection	Healthcare associated bloodstream infection	Unit 1: HR = 1.53; 95% CI = 0.39–6.07 Unit 2: HR = 0.21; 95% CI = 0.06–0.79*
Hugonnet et al., 2007b	Nurse/patient ratio	RN	Averaged in the 4 days prior to pneumonia	Early onset Ventilator-Associated Pneumonia	Not significant and not reported
Hugonnet et al., 2007b	Nurse/patient ratio	RN	Averaged in the 4 days prior to pneumonia	Late onset Ventilator Associated Pneumonia	HR = 0.42; 95% CI = 0.18–0.99*
Alonso-Echanove et al., 2003	Nurse/patient ratio	RN	Averaged across patient stay	CVC-associated blood stream infections	No estimates - reported as nonsignificant
Alonso-Echanove et al., 2003	Nurse/patient ratio	NA	Averaged across patient stay	CVC-associated blood stream infections	No estimates - reported as nonsignificant
Alonso-Echanove et al., 2003	Being cared for by float nurses for >60% of the time	RN	Averaged across patient stay	CVC-associated blood stream infections	HR = 2.75; 95% CI = 1.45–5.22*
Lower staffing					
Beltempo et al., 2017	Ratio of available to recommended nurses 0.98–0.92	RN	Three days prior to infection	Healthcare associated infection	Compared to days with ratio ≥0.98: OR = 1.39; 95% CI = 0.89–2.20
Beltempo et al., 2017	Ratio of available to recommended nurses 0.92–0.86	RN	Three days prior to infection	Healthcare associated infection	Compared to days with ratio ≥0.98: OR = 1.24; 05% CI = 0.77–2.00
Beltempo et al., 2017	Ratio of available to recommended nurses <0.86	RN	Three days prior to infection	Healthcare associated infection	Compared to days with ratio ≥0.98: OR = 1.16; 95% CI = 0.67–1.99
Shang et al., 2019	Number of shifts with low nurse hours per patient shift	RN	Cumulative sum of shifts below 80% of annual median unit over the two days prior to infection	Healthcare associated infection	No estimates - reported as nonsignificant
Shang et al., 2019	Number of shifts with low nurse hours per patient shift	RN	Any shift below 80% of annual median unit over the two days prior to infection	Healthcare associated infection	HR = 1.15; 95% CI = 1.02–1.30*
Shang et al., 2019	Number of shifts with low nurse hours per patient shift	NA	Cumulative sum of shifts below 80% of annual median unit over the two days prior to infection	Healthcare associated infection	No estimates - reported as non-significant
Shang et al., 2019	Number of shifts with low nurse hours per patient shift	NA	Any shift below 80% of annual median unit over the two days prior to infection	Healthcare associated infection	HR = 1.11; 95% CI = 1.01–1.21*

* Statistically significant at <0.05; OR = Odds Ratio; HR = Hazard Ratio; RR = Risk Ratio; CI = Confidence Interval.

exposed to risk means that these effects are, nonetheless, important. In a study that included 138,000 general medical and surgical admissions to one large hospital with approximately 800 medical surgical beds, over 3 years it was estimated that an increase of 1 registered nurse hour per patient day over could avoid 657 deaths and save over 30,000 bed days through reduced length of stay (Griffiths et al., 2018b).

When considering nursing assistants and other grades of nursing staff, the effects observed are mixed, and so a general conclusion and causal inference from these results is more challenging. It seems likely from the evidence that any causal relationship is complex, with the hints of non-linear effects (Griffiths et al., 2019; Musy et al., 2021) suggesting competing causal mechanisms from increased resource (beneficial) as opposed to substitution of registered nurses (harmful). Findings around temporary staffing, albeit from a single study, show that when small proportions of temporary registered nurses are deployed there is no evidence of harm, presumably because staffing levels are being maintained, but when high levels of temporary staff are deployed, the association with mortality changes direction.

Regarding other patient outcomes, there is more uncertainty. This is largely due to higher risk of bias in studies. Nonetheless, the direction of the relationships observed, particularly for infections, pressure ulcers and length of stay, remains compatible with a protective effect of higher nurse staffing levels. While much work has been done already about nurse-sensitive outcomes (Twigg et al., 2015; Griffiths et al., 2008; Blume et al., 2021), the current evidence does not provide a consistent and coherent overview of how nurse staffing affects other patient outcomes. Estimation of the effect of variation in nurse staffing levels on some well-established nurse-sensitive outcomes, including falls and pressure ulcers, is hampered by inadequate risk-adjustment models, poor recording, and ascertainment bias in the current research.

While there is evidence for the link between nurse staffing levels and mortality that is strong from an internal validity perspective, large-scale longitudinal studies from a larger number of hospitals are needed to increase external validity, improving both the precision and generalisability of estimates. While we did not formally assess external validity, all studies at low risk of bias were conducted in a single hospital, which limits generalisability of findings. Convincingly demonstrating external validity will be crucial to change policy and practice around nurse staffing levels (Glasgow et al., 2006; Burchett et al., 2011).

In addition, we note that staffing levels were measured in a variety of ways across studies. If approaches to measuring staffing inputs were standardised, or if raw anonymised data were provided alongside papers, more comparable estimates could be combined in meta-analyses, and the field of research and practice might be advanced (Manojlovich et al., 2011). While it is difficult to establish the clear superiority of one staffing effect measure over another, we would encourage all authors to offer analyses using measures based on those used in existing reports in addition to any de-novo measures they derive. A way forward might be using a measured staffing requirement as a reference, as used for example by Griffiths et al., 2018a, Needleman et al., 2011, but the diversity of systems for determining staffing requirements and limited evidence for validity of any approach would still make standardisation and comparison challenging (Griffiths et al., 2020). Staffing deviation standardised against unit norms (mean, median) has been used as an absolute (low staffing) and continuous (hours per patient day relative to the norm) effect measure, which could be derived for all studies irrespective of the underlying methods used to determine staffing requirements, although such methods are limited to estimating the effect of within-unit variation.

4.1. Limitations of the review

Although our search was extensive, the topic is difficult to capture precisely in structured searches. As a team we are familiar with the literature and were thus able to test the ability of strategies to identify already known studies and thus be confident about the overall sensitivity of the search, it is still possible that we missed some eligible studies. However, it seems unlikely that we would have missed a number of low-risk of bias studies sufficient to fundamentally change our conclusions or the general picture of the literature that we have presented. Our review was conducted according to a pre-defined protocol, but the diversity of designs we encountered meant that a small number of studies that did not meet our inclusion criteria, but which still explored trends over time were excluded because the measured variation over time did not precede the reported outcome. For example, one study explored staffing over several years and associated annual staffing with annual outcomes (He et al., 2016). While these studies may be informative, they were small in number and their inclusion would not be likely to substantively change our conclusions.

5. Conclusions

By focussing on longitudinal studies in this review we have addressed a significant critique of the body of research linking nurse staffing to patient outcomes. While causal conclusions from associations observed in cross-sectional studies may be dismissed as spurious, we have revealed a substantial body of longitudinal evidence, which is harder to dismiss. The evidence reviewed remains observational and potentially subject to bias in estimating effects, but a small number of studies at low risk of bias provide powerful support for a causal interpretation of the findings. Higher registered nurse staffing levels reduce the risk of patient death in acute care settings, although the limited number of studies at low risk of bias makes it difficult to provide generalisable estimates of effect. In contrast with clear evidence for benefits from increased registered nurse staffing, mixed evidence showing both benefits and harms from adding other grades of staff to the nursing team is harder to interpret. While it seems possible that such staff may make some contribution to patient safety, the evidence cannot be used to support substitution of registered nurses by other grades of staff.

Funding

This study is funded by the NIHR [Health and Social Care Delivery Research (NIHR128056) and Applied Research Collaboration Wessex]. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Declaration of Competing Interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijnurstu.2022.104311.

References

Al-Abdwani, R., Williams, C.B., Dunn, C., Macartney, J., Wollny, K., Frndova, H., Chin, N., Stephens, D., Parshuram, C.S., 2018. Incidence, outcomes and outcome prediction of unplanned extubation in critically ill children: an 11year experience. J. Crit. Care 44, 368–375.

Alonso-Echanove, J., Edwards, J.R., Richards, M.J., Brennan, P., Venezia, R.A., Keen, J., Ashline, V., Kirkland, K., Chou, E., Hupert, M., Veeder, A.V., Speas, J., Kaye, J., Sharma, K., Martin, A., Moroz, V.D., Gaynes, R.P., 2003. Effect of nurse staffing and antimicrobial-impregnated central venous catheters on the risk for bloodstream infections in intensive care units. Infect. Control Hosp. Epidemiol. 24, 916–925.

Ambrosi, E., de Togni, S., Guarnier, A., Barelli, P., Zambiasi, P., Allegrini, E., Bazoli, L., Casson, P., Marin, M., Padovan, M., Picogna, M., Taddia, P., Salmaso, D., Chiari, P., Frison, T., Marognolli, O., Canzan, F., Saiani, L., Palese, A., 2017. In-hospital elderly mortality and associated factors in 12 Italian acute medical units: findings from an exploratory longitudinal study. Aging Clin. Exp. Res. 29, 517–527.

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R., 2010. On making causal claims: a review and recommendations. Leadersh. Q. 21, 1086–1120.

Ball, J.E., Bruyneel, L., Aiken, L.H., Sermeus, W., Sloane, D.M., Rafferty, A.M., Lindqvist, R., Tishelman, C., Griffiths, P., Consortium, R.N.C., 2018. Post-operative mortality, missed care and nurse staffing in nine countries: a cross-sectional study. Int. J. Nurs. Stud. 78, 10–15.

Beech, J., Bottery, S., Charlesworth, A., Evans, H., Gershlick, B., Hemmings, N., Imison, C., Kahtan, P., McKenna, H., Murray, R., 2019. Closing the gap. Key areas for action on the health and care workforce. The Health Foundation/Nuffield Trust/The King's Fund, London.

Beltempo, M., Blais, R., Lacroix, G., Cabot, M., Piedboeuf, B., 2017. Association of Nursing Overtime, nurse staffing, and unit occupancy with health care-associated infections in the NICU. Am. J. Perinatol. 34, 996–1002.

Beltempo, M., Lacroix, G., Cabot, M., Blais, R., Piedboeuf, B., 2018. Association of nursing overtime, nurse staffing and unit occupancy with medical incidents and outcomes of very preterm infants. J. Perinatol. 38, 175–180.

Blume, K.S., Dietermann, K., Kirchner-Heklau, U., Winter, V., Fleischer, S., Kreidl, L.M., Meyer, G., Schreyogg, J., 2021. Staffing levels and nursing-sensitive patient outcomes: umbrella review and qualitative study. Health Serv Res 56 (5), 885–907.

Buchan, J., Ball, J., Shembavnekar, N., Charlesworth, A., 2020. Building the NHS Nursing Workforce in England.

Burchett, H., Umoquit, M., Dobrow, M., 2011. How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. J. Health Serv. Res. Policy 16, 238–244.

Callaghan, L.A., Cartwright, D.W., O'Rourke, P., Davies, M.W., 2003. Infant to staff ratios and risk of mortality in very low birthweight infants. Arch. Dis. Child. Fetal Neonatal Ed. 88, F94–F97.

Cimiotti, J.P., 2004. Nurse Staffing and Healthcare-associated Infections in the Neonatal ICU. Columbia University D.N.Sc.

Cimiotti, J.P., Haas, J., Saiman, L., Larson, E.L., 2006. Impact of staffing on bloodstream infections in the neonatal intensive care unit. Arch. Pediatr. Adolesc. Med. 160, 832–836.

Dall'Ora, C., Maruotti, A., Griffiths, P., 2020. Temporary staffing and patient death in acute care hospitals: a retrospective longitudinal study. J. Nurs. Scholarsh. 52, 210–216.

Driscoll, A., Grant, M.J., Carroll, D., Dalton, S., Deaton, C., Jones, I., Lehwaldt, D., McKee, G., Munyombwe, T., Astin, F., 2018. The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: a systematic review and meta-analysis. Eur. J. Cardiovasc. Nurs. 17, 6–22.

Fogg, C., Bridges, J., Meredith, P., Spice, C., Field, L., Culliford, D., Griffiths, P., 2021. The association between ward staffing levels, mortality and hospital readmission in older hospitalised adults, according to presence of cognitive impairment: a retrospective cohort study. Age Ageing 50, 431–439.

Glasgow, R.E., Green, L.W., Klesges, L.M., Abrams, D.B., Fisher, E.B., Goldstein, M.G., Hayman, L.L., Ockene, J.K., Orleans, C.T., 2006. External validity: we need to do more. Ann. Behav. Med. 31, 105–108.

Griffiths, P., Ball, J., Bloor, K., Böhning, D., Briggs, J., Dall'ora, C., Iongh, A.D., Jones, J., Kovacs, C., Maruotti, A., Meredith, P., Prytherch, D., Saucedo, A.R., Redfern, O., Schmidt, P., Sinden, N., Smith, G., 2018. Nurse staffing levels, missed vital signs and mortality in hospitals: retrospective longitudinal observational study. Health Serv. Deliv. Res. J. 6.

Griffiths, P., Ball, J., Bloor, K., Böhning, D., Briggs, J., Dall'ora, C., Iongh, A.D., Jones, J., Kovacs, C., Maruotti, A., Meredith, P., Prytherch, D., Saucedo, A.R., Redfern, O., Schmidt, P., Sinden, N., Smith, G., 2018. Nurse staffing levels, missed vital signs and mortality in hospitals: retrospective longitudinal observational study. NIHR J. Libr. 6 (38) Southampton (UK).

Griffiths, P., Ball, J., Drennan, J., Dall'ora, C., Jones, J., Maruotti, A., Pope, C., Recio Saucedo, A., Simon, M., 2016. Nurse staffing and patient outcomes: Strengths and limitations of the evidence to inform policy and practice. A review and discussion paper based on evidence reviewed for the National Institute for Health and Care Excellence Safe Staffing guideline development. Int. J. Nurs. Stud. 63, 213–225.

Griffiths, P., Jones, S., Maben, J., Murrels, T., 2008. State of the Art Metrics for Nursing: A Rapid Appraisal. National Nursing Research Unit.

Griffiths, P., Maruotti, A., Recio Saucedo, A., Redfern, O.C., Ball, J.E., Briggs, J., Dall'ora, C., Schmidt, P.E., Smith, G.B., 2019. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. BMJ Qual. Saf. 28, 609–617.

Griffiths, P., Saville, C., Ball, J., Jones, J., Pattison, N., Monks, T., 2020. Nursing workload, nurse staffing methodologies and tools: a systematic scoping review and discussion. Int. J. Nurs. Stud. 103, 103487.

Haegdorens, F., van Bogaert, P., de Meester, K., Monsieurs, K.G., 2019. The impact of nurse staffing levels and nurse's education on patient mortality in medical and surgical wards: an observational multicentre study. BMC Health Serv. Res. 19, 864.

Hamilton, K.E., Redshaw, M.E.S., Tarnow-Mordi, W., 2007. Nurse staffing in relation to risk-adjusted mortality in neonatal care. Arch. Dis. Child. Fetal Neonatal Ed. 92, F99–F103.

He, J., Almenoff, P.L., Keighley, J., Li, Y.F., 2013. Impact of patient-level risk adjustment on the findings about nurse staffing and 30-day mortality in veterans affairs acute care hospitals. Nurs. Res. 62, 226–232.

He, J., Staggs, V.S., Bergquist-Beringer, S., Dunton, N., 2016. Nurse staffing and patient outcomes: a longitudinal study on trend and seasonality. BMC Nurs. 15, 60.

Hernán, M.A., 2018. The C-word: scientific euphemisms do not improve causal inference from observational data. Am. J. Public Health 108, 616–619.

Higgins, J., Savović, J., Pae, M.J., Elbers, R.G., Jac, S., 2021. Chapter 8: Assessing risk of bias in a randomized trial. In: JPT, Higgins, Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Va, W. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 Cochrane.

Hill, A.B., 1965. The environment and disease: association or causation? Proc. R. Soc. Med. 58, 295–300.

Hugonnet, S., Chevrolet, J., Pittet, D., 2007a. The effect of workload on infection risk in critically ill patients. Crit. Care Med. 35, 76–81.

Hugonnet, S., Uckay, I., Pittet, D., 2007b. Staffing level: a determinant of late-onset ventilator-associated pneumonia. Crit. Care 11, R80.

Jansson, M.M., Syrjälä, H.P., Ala-Kokko, T.I., 2019. Association of nurse staffing and nursing workload with ventilator-associated pneumonia and mortality: a prospective, single-center cohort study. J. Hosp. Infect. 101, 257–263.

Kane, R.L., Shamliyan, T.A., Mueller, C., Duval, S., Wilt, T.J., 2007. The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. Med. Care 45, 1195–1204.

Manojlovich, M., Sidani, S., Covell, C.L., Antonakos, C.L., 2011. Nurse dose: linking staffing variables to adverse patient outcomes. Nurs. Res. 60, 214–220.

Mark, B.A., Belyea, M., 2009. Nurse staffing and medication errors: cross-sectional or longitudinal relationships? Research in Nursing & Health 32, 18–30.

McHugh, M.D., Aiken, L.H., Sloane, D.M., Windsor, C., Douglas, C., Yates, P., 2021. Effects of nurse-to-patient ratio legislation on nurse staffing and patient mortality, readmissions, and length of stay: a prospective study in a panel of hospitals. Lancet 397, 1905–1913.

Musy, S.N., Endrich, O., Leichtle, A.B., Griffiths, P., Nakas, C.T., Simon, M., 2020. Longitudinal study of the variation in patient turnover and patient-to-nurse ratio: descriptive analysis of a Swiss university hospital. J. Med. Internet Res. 22 N.PAG-N.PAG.

Musy, S.N., Endrich, O., Leichtle, A.B., Griffiths, P., Nakas, C.T., Simon, M., 2021. The association between nurse staffing and inpatient mortality: a shift-level retrospective longitudinal study. Int. J. Nurs. Stud. 103950.

Needleman, J., Buerhaus, P., Pankratz, V.S., Leibson, C.L., Stevens, S.R., Harris, M., 2011. Nurse staffing and inpatient hospital mortality. N. Engl. J. Med. 364, 1037–1045.

Needleman, J., Jianfang, L., Jinjing, S., Larson, E.L., Stone, P.W., 2020. Association of registered nurse and nursing support staffing with inpatient hospital mortality. BMJ Qual. Saf. 29, 10–18.

Palese, A., Gonella, S., Moreale, R., Guarnier, A., Barelli, P., Zambiasi, P., Allegrini, E., Bazoli, L., Casson, P., Marin, M., Padovan, M., Picogna, M., Taddia, T., Salmaso, D., Chiari, P., Frison, T., Marognolli, O., Benaglio, C., Canzan, F., Ambrosi, E., Saiani, L., Group, E., 2016. Hospital-acquired functional decline in older patients cared for in acute medical wards and predictors: findings from a multicentre longitudinal study. Geriatr. Nurs. 37, 192–199.

Patrician, P.A., McCarthy, M.S., Swiger, P., Raju, D., Breckenridge-Sproat, S., Su, X., Randall, K.H., Loan, L.A., 2017. Association of Temporal Variations in staffing with hospital-acquired pressure injury in military hospitals. Res Nurs Health 40, 111–119.

Redfern, O.C., Griffiths, P., Maruotti, A., Recio Saucedo, A., Smith, G.B., Missed Care Study G, 2019. The association between nurse staffing levels and the timeliness of vital signs monitoring: a retrospective observational study in the UK. BMJ Open 9, e032157.

Rochefort, C.M., Beauchamp, M.-E., Audet, L.-A., Abrahamowicz, M., Bourgault, P., 2020. Associations of 4 nurse staffing practices with hospital mortality. Med. Care 58, 912–918.

Rothman, K.J., Greenland, S., 2005. Causation and causal inference in epidemiology. Am. J. Public Health 95 (Suppl. 1), S144–S150.

Shang, J., Needleman, J., Liu, J., Larson, E., Stone, P.W., 2019. Nurse staffing and healthcare-associated infection, unit-level analysis. J. Nurs. Adm. 49, 260–265.

Shekelle, P.G., 2013. Nurse–patient ratios as a patient safety strategy: a systematic review. Ann. Intern. Med. 158, 404–409.

Shin, S., Park, J.H., Bae, S.H., 2019. Nurse staffing and hospital-acquired conditions: a systematic review. J. Clin. Nurs. 28, 4264–4275.

Smith, G.B., Redfern, O., Maruotti, A., Recio-Saucedo, A., Griffiths, P., 2020. The association between nurse staffing levels and a failure to respond to patients with deranged physiology: a retrospective observational study in the UK. Resuscitation 149, 202–208.

Sterne, J.A., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I., Carpenter, J.R., Chan, A.-W., Churchill, R., Deeks, J.J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y.K., Pigott, T.D., Ramsay, C.R., Regidor, D., Rothstein, H.R., Sandhu, L., Santaguida, P.L., Schünemann, H.J., Shea, B., Shrier, I., Tugwell, P., Turner, L., Valentine, J.C., Waddington, H., Waters, E., Wells, G.A., Whiting, P.F., Higgins, J.P., 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ (Clin. Res. Ed.) 355 i4919-i4919.

Tschannen, D.J., 2005. Organizational Structure, Process, and Outcome: The Effects of Nurse Staffing and Nurse-Physician Collaboration on Patient Length of Stay. University of Michigan Ph.D.

Twigg, D.E., Gelder, L., Myers, H., 2015. The impact of understaffed shifts on nurse-sensitive outcomes. J. Adv. Nurs. 71, 1564–1572.

Twigg, D.E., Kutzer, Y., Jacob, E., Seaman, K., 2019. A quantitative systematic review of the association between nurse skill mix and nursing-sensitive patient outcomes in the acute care setting. J. Adv. Nurs. 75, 3404–3423.

Twigg, D.E., Myers, H., Duffield, C., Pugh, J.D., Gelder, L., Roche, M., 2016. The impact of adding assistants in nursing to acute care hospital ward nurse staffing on adverse patient outcomes: an analysis of administrative health data. Int. J. Nurs. Stud. 63, 189–200.

World Health Organization, 2016. Global Strategic Directions for Strengthening Nursing and Midwifery 2016-2020.

Yakusheva, O., Rambur, B., Buerhaus, P.I., 2020. Value-informed nursing practice can help reset the hospital-nurse relationship. JAMA Health Forum 1, e200931.