< **e-component file 1. Figures 1-3** >

1 “nurse staffing” OR "nursing time" OR "nurse time" OR “nursing hours” OR “nurse hours” OR “hours per patient day” OR "staffing ratios" OR “nurse-patient ratio” OR “patient-nurse ratio” OR “patient to nurse ratio” OR “nurse to patient ratio” OR "bed to nurse ratio" OR "nurse to bed ratio" OR "bed nurse ratio" OR "nurse bed ratio" OR "rationing nursing" OR “number of nurses” OR “nurse numbers” OR staffing OR understaffing OR “under staffing” OR “labour force” OR “labor force” OR “healthcare force” OR workload

2. “critical care” OR "intensive care" OR "surgical care" OR "high dependency" OR "critically ill" OR "intensive therapy" OR ICU

3. 1 AND 2

**SI Figure 1: Example of MEDLINE Search Strategy**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Data Extraction: QUANTITATIVE** | | **Title:** | | | **Reviewer:** | |
| **1. Study details** | **2. Setting and Population** | **3. Sample** | **4. Risk adjustment** | **6. Results: which outcomes of staffing level variation were analysed?** | **7. Is impact of staffing significant? Significant or not, direction of effect? size of effect and 95% CIs?** | **8. Notes/comments** |
| 1a. First author (Year): | 2a. Country: | 3a. Staff groups studied: | 4a. What were the patient and/or nurse, risk adjustments: | 6a. Patient outcomes: | 7a. Patient outcomes: |  |
| 1b. Study Aim: | 2b. Setting(s): | 3b. Staffing level variables measured.  3bi Description of the variable (e.g. number of nurses per 8 hour shift):  3bii Units: (e.g. N:P, number of nurses)  3biii Was this measured specifically for the study or was this treated as a general characteristic of the unit?  3biv If measured specifically for the study, was this measured just once or for e.g., per month, per day, per shift, per patient? |
| 1c. Study design: | 2c. Source population: | 3c. Sample size: number of RN level nurses (out of total number of staff, if part of a wider study): | 4b. What were the unit/hospital risk adjustments: | 6b. Nurse outcomes: | 7b. Nurse outcomes: |
| 1d. Internal validity:  2 1 0 | 2d. Selection procedure | 3d. Sample size: Hospitals | **5.** **Analysis** | 6c. Care quality outcomes: Process/ organisational outcomes: | 7c. Care quality outcomes: Process/ organisational outcomes: |
| 5a. How was link between staffing levels and outcome(s) statistically evaluated *(e.g. logistic regression, correlation etc.)* |
| 1e. External validity:  2 1 0 | 3e. Sample size: ICU/critical care units (out of total number of units if part of a wider study) | 6d. Family outcomes: | 7d. Family outcomes: |

**SI Figure 2: Data Extraction Framework**

|  |  |  |  |
| --- | --- | --- | --- |
| **Quality Appraisal** | | | |
| **Title:**  **First author (year):** | | **Reviewer:** | |
| **Key research question/aim:** | | **Validity scores** | |
| **Internal** | **External** |
| Critical Appraisal breakdown: | Rough guide to allocating 2 (strong methodology), 1 (moderate) and 0 (weak) | 2 Strong, 1 Moderate, 0 Weak | |
| **1. Is the setting applicable to an international definition of intensive care units?** | Marshall et al., (2017) “An intensive care unit (ICU) is an organized system for the provision of care to critically ill patients that provides intensive and specialized medical and nursing care, an enhanced capacity for monitoring, and multiple modalities of physiologic organ support to sustain life during a period of acute organ system insufficiency.” **Yes 2** **uncertain/ mixed 1** **no = 0 and exclude** |  |  |
| **2. Design: how are potential associations between exposure to a change in staffing level and patient, family, staff, organisational or care quality outcome examined?** | **cross sectional** is **weak 0** (exposure and outcome assessed at the same time on an individual level); **retrospective** is **moderate 1** (e.g. group sharing the same target outcome is identified and then retrospectively checked for the likelihood of having been exposed to low/high staffing); **prospective** study allowing for cause / effect is **strong 2** (i.e. exposure precedes outcome) |  |  |
| **3. Is the eligible population or area representative of the source population or area?** | Consider whether hospitals potentially included in the study are representative of acute general hospitals in that country / state (**1**)  Were the wards/ staff / patients eligible to be included representative of ICU/critical care units/ RNs/ICU, critical care patients (**another 1**) |  |  |
| **4. Do the selected participants or areas represent the eligible population or area?** | What % of selected hospitals agreed to participate (60% plus, **1**)  What % of eligible individuals (staff / patients) participated (60% plus, **1**)  Were any data-sets derived from administrative systems complete? Were the inclusion or exclusion criteria explicit and appropriate? (**1** for either if not already two) |  |  |
| **5. Were the outcome measures reliable?** | Were main patient outcome measures subjective or objective **(2 for objective)**  For **subjective**: How reliable were outcome measures (e.g. inter- or intra-rater reliability scores?) **1**  Was there any indication that measures had been validated (e.g. validated against a gold standard measure or assessed for content? **1** |  |  |
| **6. Were the outcome measurements complete?** | Were all or most of the study participants who met the defined study outcome definitions likely to have been identified? **(2 for definitely … e.g. patient mortality**, **1 for outcomes collected using clearly defined methods but where some may have been missed**, **0 for info. abstracted from e.g. discharge abstracts**) |  |  |
| **7. Was the study sufficiently powered to detect an effect (if one exists)?** | Were there sufficient units / hospitals / wards to give variation and enough patients to detect effects? Generally, look at effect sizes and units of measure; a large clinically important effect size which is not significant suggests the study is underpowered. Some guidance:  Large multi-hospital (20+) studies (state / national / international) with administrative data **2**  Smaller studies / single hospital with large numbers of patients (000,000’s) **1,** Other **0** |  |  |
| **8. How well were likely confounding factors identified and controlled?** | Risk adjustment for main outcomes. Was there patient/staff level risk adjustment for e.g. for AGE, DIAGNOSIS, COMORBIDITY, YEARS EXPERIENCE **(2, 1 moderate, 0 none**) |  |  |
| **9. Were the analytical methods appropriate?** | Was there adjustment for clustering of data within wards / hospitals? (**1)** Where relevant was there control for ward / hospital characteristics (**1**) |  |  |
| **10. Was the precision of association given or calculable? Is association meaningful?** | Were confidence intervals or p values for effect estimates given or possible to calculate? (**1**) Were CIs wide or were they sufficiently precise to aid decision-making? (**1**) If precision is lacking, is this because the study is underpowered? (**0**) |  |  |
|  |  | **Overall internal validity (bias) 2 1 0** | **Overall external validity**  **2 1 0** |
| **11. Are the study results internally valid (i.e. unbiased)?** | How well did the study minimise sources of bias (i.e. adjusting for potential confounders)?  Were there significant flaws in the study design? |  |  |
| **12. Are the findings generalisable to the source population (i.e. externally valid)?** | Are there sufficient details given about the study to determine if the findings are generalisable to the source population? Consider: participants, interventions and comparisons, outcomes, resource and policy implications. |  |  |

**SI Figure 3: Critical Appraisal Framework.**